PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/85, C07K 14/47, 16/18, C12Q 1/68

A2

(11) International Publication Number: WO 99/32643

(43) International Publication Date: 1 July 1999 (01.07.99)

(21) International Application Number: PCT/EP98/08543

(22) International Filing Date: 17 December 1998 (17.12.98)

(30) Priority Data:
9726804.9 18 December 1997 (18.12.97) GB

(71) Applicant *(for all designated States except US)*: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE [BE/BE]; Rijvisschestraat 118 Bus 1, B-9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VAN BROECKHOVEN, Christine [BE/BE]; Koning Albertlei 15, B-2650 Edegem (BE). RAEYMAEKERS, Peter [BE/BE]; Kardinaal Cardijnlaan 104, B-2547 Lint (BE). DEL-FAVERO, Jurgen [BE/BE]; Oorbeeksesteenweg 149, B-3300 Tienen (BE).

(74) Agents: BALDOCK, Sharon, Claire et al.; Boult Wade Tennant, 27 Furnival Street, London EC4A 1PQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MOOD DISORDER GENE

(57) Abstract

The present invention comprises the use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders. The invention also provides methods for determining the susceptibility of an individual to mood disorders or related disorders, comprising analysing a DNA sample for the presence of a trinucleotide repeat expansion in the above region. Polynucleotide sequences useful for detecting the presence of such trinucleotide repeat expansions are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

MOOD DISORDER GENE

The invention is concerned with the determination of genetic factors associated with psychiatric health with particular reference to a human gene or genes which contributes to or is responsible for the manifestation of a mood disorder or a related disorder in affected individuals. In particular, although not exclusively, the invention provides a method of identifying and characterising such a gene or genes from human chromosome 18, as well as genes so identified and their expression products. The invention is also concerned with methods of determining the genetic susceptibility of an individual to a mood disorder or related disorder. By mood disorders or related disorders is meant the following disorders as defined in the Diagnostic and Statistical Manual of Mental Disorders, version 4 (DSM-IV) taxonomy (DSM-IV codes in parenthesis):- mood disorders (296.XX, 300.4, 311, 301.13, 295.70), schizophrenia and related disorders (295.XX, 297.1,298.8, 297.3, 298.9), anxiety disorders (300.XX, 309.81,308.3), adjustment disorders (309.XX) and personality disorders (codes 301.XX).

The methods of the invention are particularly exemplified in relation to genetic factors associated with a family of mood disorders known as Bipolar (BP) spectrum disorders.

Bipolar disorder (BP) is a severe psychiatric condition that is characterized by disturbances in mood, ranging from an extreme state of elation (mania) to a severe state of dysphoria (depression). Two types of bipolar illness have been described: type I BP illness (BPI) is characterized by major depressive episodes alternated with phases of mania, and type II

BP illness (BPII), characterized by major depressive episodes alternating with phases of hypomania. Relatives of BP probands have an increased risk for BP, unipolar disorder (patients only experiencing depressive episodes; UP), cyclothymia (minor depression and hypomania episodes; CY) as well as for schizoaffective disorders of the manic (SAm) and depressive (SAd) type. Based on these observations BP, CY, UP and SA are classified as BP spectrum disorders. The involvement of genetic factors in the etiology of BP spectrum disorders was suggested by family, twin and adoption studies (Tsuang and Faraone (1990), The Genetics of Mood Disorders, Baltimore, The John Hopkins University Press). However, the exact pattern of transmission is unknown. In some studies, complex segregation analysis supports the existence of a single major locus for BP (Spence *et al.* (1995), Am J. Med. Genet (Neuropsych. Genet.) 60 pp 370-376). Other researchers propose a liability-threshold-model, in which the liability to develop the disorder results from the additive combination of multiple genetic and environmental effects (McGuffin *et al.* (1994), Affective Disorders; Seminars in Psychiatric Genetics Gaskell, London pp 110-127).

Due to the complex mode of inheritance, parametric and nonparametric linkage strategies are applied in families in which BP disorder appears to be transmitted in a Mendelian fashion. Early linkage findings on chromosomes 11p15 (Egeland *et al.* (1987), Nature **325** pp 783-787) and Xq27-q28 (Mendlewicz *et al.* (1987) The Lancet **1** pp 1230 -1232; Baron *et al.* (1987) Nature **326** pp 289-292) have been controversial and could initially not be replicated (Kelsoe *et al.* (1989) Nature **242** pp 238-243; Baron *et al.* (1993) Nature Genet **3** pp 49-55). With the development of a

human genetic map saturated with highly polymorphic markers and the continuous development of data analysis techniques, numerous new linkage searches were started. In several studies, evidence or suggestive evidence for linkage to particular regions on chromosomes 4, 12, 18, 21 and X was found (Blackwood *et al.* (1996) Nature Genetics **12** pp 427-430, Craddock *et al.* (1994) Brit J. Psychiatry **164** pp 355-358, Berrettini *et al.* (1994), Proc Natl Acad Sci USA **91** pp 5918-5921, Straub *et al.* (1994) Nature Genetics **8** pp 291-296 and Pekkarinen *et al.* (1995) Genome Research **5** pp 105-115). In order to test the validity of the reported linkage results, these findings have to be replicated in other, independent studies.

Recently, linkage of bipolar disorder to the pericentromeric region on chromosome 18 was reported (Berrettini *et al.* 1994). Also a ring chromosome 18 with break-points and deleted regions at 18pter-p11 and 18q23-qter was reported in three unrelated patients with BP illness or related syndromes (Craddock *et al.* 1994). The chromosome 18p linkage was replicated by Stine *et al.* (1995) Am J Hum Genet **57** pp 1384-1394, who also reported suggestive evidence for a locus on 18q21.2-q21.32 in the same study. Interestingly, Stine *et al.* observed a parent-of-origin effect: the evidence of linkage was the strongest in the paternal pedigrees, in which the proband's father or one of the proband's father's sibs is affected.

In an independent replication study, the present inventors tested linkage with chromosome 18 markers in 10 Belgian families with a bipolar proband. To localize causative genes the linkage analysis or likelihood method was used in these families. This

method studies within a family the segregation of a defined disease phenotype with that of polymorphic genetic markers distributed in the human genome. The likelihood ratio of observing cosegregation of the disease and a genetic marker under linkage versus no linkage is calculated and the log of this ratio or the log of the odds is the LOD score statistic z. A LOD score of 3 (or likelihood ratio of 1000 or greater) is taken as significant statistical evidence for linkage. In the inventors' study no evidence for linkage to the pericentromeric regions was found, but in one of the families, MAD31, a Belgian family of a BPII proband, suggestive linkage was found with markers located at 18q21.33-q23 (De bruyn *et al.* (1996) Biol Psychiatry **39** pp 679-688). Multipoint linkage analysis gave the highest LOD score in the interval between STR (Short Tandem Repeats) polymorphisms D18S51 and D18S61, with a maximum multipoint LOD score of +1.34. Simulation studies indicated that this LOD score is within the range of what can be expected for a linked marker given the information available in the family. Likewise, an affected sib-pair analysis also rejected the null-hypothesis of nonlinkage for several of the markers tested. Two other groups also found evidence for linkage of bipolar disorder to 18q (Freimer *et al.* (1996) Nature Genetics **12** pp 436-441, Coon *et al.* (1996) Biol Psychiatry **39** pp 689 to 696). Although the candidate regions in the different studies do not entirely overlap, they all suggest the presence of a susceptibility locus at 18q21-q23.

The inventors have now carried out further investigations into the 18q chromosomal region in family MAD31. By analysis of cosegregation of bipolar disease in MAD31 with twelve STR polymorphic markers previously located between the aforementioned markers

D18S51 and D18S61 and subsequent LOD score analysis as described above, the inventors have further refined the candidate region of chromosome 18 in which a gene associated with mood disorders such as bipolar spectrum disorders may be located and have constructed a physical map. The region in question may thus be used to locate, isolate and sequence a gene or genes which influences psychiatric health and mood.

The inventors have also constructed a YAC (yeast artificial chromosome) contig map of the candidate region to determine the relative order of the twelve STR markers mapped by the cosegregational analysis and they have identified seven clones from the YAC library incorporating the candidate region.

A number of procedures can be applied to the identified YAC clones and, where applicable, to the DNA of an individual afflicted with a mood disorder as defined herein, in the process of identifying and characterising the relevant gene or genes. For example, the inventors have used YAC clones spanning the region of interest in chromosome 18 to identify by CAG or CTG fragmentation novel genes that are allegedly involved in the manifestation of mood disorders or related disorders.

Other procedures can also be applied to the said YAC clones to identify candidate genes as discussed below.

Once candidate genes have been identified it is possible to assess the susceptibility of an individual to a mood disorder or related disorder by detecting the presence of a polymorphism associated with a mood disorder or related disorder in such genes.

Accordingly, in a first aspect the present invention comprises the use of an 8.9 cM region of

human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders as defined above. As will be described below, the present inventors have identified this candidate region of chromosome 18q for such a gene, by analysis of cosegregation of bipolar disease in family MAD31 with 12 STR polymorphic markers previously located between D18S51 and D18S61 and subsequent LOD score analysis.

In a second aspect the invention comprises the use of a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61 for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with mood disorders or related disorders as defined above. D18S60 is close to D18S51 so the particular YAC clones for use are those which have an artificial chromosome spanning the candidate region of human chromosome 18q between polymorphic markers D18S51 and D18S61 as identified by the present inventors in their earlier paper (De bruyn *et al.* (1996)).

Particular YACs covering the candidate region which may be used in accordance with the present invention are 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752-g-8 and 717_d_3, preferred ones being 961_h_9, 766_f_12 and 907_e_1 since these have the minimum tiling path across the candidate region. Suitable YAC clones for use are those having an artificial chromosome spanning the refined candidate region between D18S68 and D18S979.

There are a number of methods which can be applied to the candidate regions of chromosome 18q as defined above, whether or not present in a YAC, to identify a candidate gene or genes associated with mood disorders or related disorders. For example, it has previously been demonstrated that an apparent association exists between the presence of trinucleotide repeat expansions (TRE) in the human genome and the phenomenon of anticipation of mood disorders (Lindblad *et al.* (1995), Neurobiology of Disease **2**: 55-62 and O'Donovan *et al.* (1995), Nature Genetics **10**: 380-381).

Accordingly, in a third aspect the present invention comprises a method of identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder as defined herein which comprises detecting nucleotide triplet repeats in the region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

An alternative method of identifying said gene or genes comprises fragmenting a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61, for example one or more of the seven aforementioned YAC clones, and detecting any nucleotide triplet repeats in said fragments. Nucleic acid probes comprising at least 5 and preferably at least 10 CTG and/or CAG triplet repeats are a suitable means of detection when appropriately labelled. Trinucleotide repeats may also be determined using the known RED (repeat expansion detection) system (Shalling et al.(1993), Nature Genetics 4 pp 135-139).

In a fourth embodiment the invention comprises a

method of identifying at least one gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder and which is present in a YAC clone spanning the region of human chromosome 18q between polymorphic markers D18S60 and D18S61, the method comprising the step of detecting the expression product of a gene incorporating nucleotide triplet repeats by use of an antibody capable of recognising a protein with an amino acid sequence comprising a string of at least 8, but preferably at least 12, continuous glutamine residues. Such a method may be implemented by subcloning YAC DNA, for example from the seven aforementioned YAC clones, into a human DNA expression library. A preferred means of detecting the relevant expression product is by use of a monoclonal antibody, in particular mAB 1C2, the preparation and properties of which are described in International Patent Application Publication No WO 97/17445.

As will be described in detail below, in order to identify candidate genes containing triplet repeats, the inventors have carried out direct CAG or CTG fragmentation of YACs 961_h_9, 766_f_12 and 907_e_1, comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61, and have identified a number of sequences containing CAG or CTG repeats, whose abnormal expansion may be involved in genetic susceptibility to a mood disorder or related disorder.

Accordingly, in a fifth aspect, the invention provides a nucleic acid comprising the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a.

In a further aspect, the invention provides a protein comprising an amino acid sequence encoded by the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a.

In yet a further aspect the invention provides a mutated nucleic acid comprising a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats.

Also provided by the invention is a mutated protein comprising an amino acid sequence encoded by a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats.

It is to be understood that the invention also contemplates nucleotide sequences having at least 75% and preferably at least 80% homology with any of the sequences described above and having functional identity with any of said sequences. The homology is calculated as described by Altschul *et al.* (1997) Nucleic Acids Res. **25**: 3389-3402, Karlin *et al.* (1990) Proc Natl Acad Sci USA **87**: 2264-68 and Karlin *et al.* (1993) Proc Natl Acad Sci USA **90**: 5873-5877. Also contemplated are amino acid sequences which differ from the above described sequences only in conservative amino acid changes. Suitable changes are well known to those skilled in the art.

Knowledge of the sequences described above can be used to design assays to determine the genetic susceptibility of an individual to a mood disorder or

related disorder.

Accordingly, in a further aspect the invention provides a method for determining the susceptibility of an individual to a mood disorder or related disorder which comprises the steps of:

a) obtaining a DNA sample from said individual;

b) providing primers suitable for the amplification of a nucleotide sequence comprised in the sequence shown in any one of Figures 15a, 16a, 17a or 18a said primers flanking the trinucleotide repeats comprised in said sequence;

c) applying said primers to the said DNA sample and carrying out an amplification reaction;

d) carrying out the same amplification reaction on a DNA sample from a control individual; and

e) comparing the results of the amplification reaction for the said individual and for the said control individual;

wherein the presence of an amplified fragment from said individual which is bigger in size from that of said control individual is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.
By control individual is meant an individual who is not affected by a mood disorder or related disorder and does not have a family history of mood disorders or related disorders.

Preferable primers to use in this method are those shown in Figure 15b, 16b, 17b or 18b but other suitable primers may be utilised.

In a further aspect the invention provides a method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises the steps of :

a) obtaining a protein sample from said individual; and

b) detecting the presence of a protein comprising an amino acid sequence encoded by a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats

wherein the presence of said protein is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

Preferably, the foresaid protein is detected by utilising an antibody that is capable of recognising a string of at least 8 continuous glutamines as, for example, the mAB 1C2 antibody.

The nucleic acids molecules according to the invention may be advantageously included in an expression vector, which may be introduced into a host cell of prokaryotic or eukaryotic origin. Suitable expression vectors include plasmids, which may be used to express foreign DNA in bacterial or eukaryotic host cells, viral vectors, yeast artificial chromosomes or mammalian artificial chromosomes. The vector may be

transfected or transformed into host cells using suitable methods known in the art such as, for example, electroporation, microinjection, infection, lipoinfection and direct uptake. Such methods are described in more detail, for example, by Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", 2nd ed. (1989) and by Ausbel *et al.* "Current Protocols in Molecular Biology", (1994).

Also provided by the invention is a host cell, tissue or organism comprising the expression vector according to the invention. The invention further provides a transgenic host cell, tissue or organism comprising a transgene capable of encoding the proteins of the invention, which may comprise a genomic DNA or a cDNA. The transgene may be present in the trangenic host cell, tissue or organism either stably integrated into the genome or in an extra chromosomal state.

A nucleic acid molecule comprising a nucleotide sequence shown in any one of Figures 15a, 16a, 17a or 18a as well as the protein encoded by it may be therapeutically used in the treatment of mood disorders or related disorders in patients which present a trinucleotide repeat expansion (TRE) in at least one of the foresaid sequences.

Accordingly, in another of its aspects the invention provides the above described nucleic acid molecules and proteins for use as medicaments for the treatment of individuals with a mood disorder or related disorder. Preferably, the nucleic acid or the protein is present in an appropriate carrier or delivery vehicle. As an example, the nucleic acid inserted into a vector, for example a plasmid or a

viral vector, may be transfected into a mammalian cell such as a somatic cell or a mammalian germ line cell, as described above. The cell to be transfected can be present in a biological sample obtained from the patient, for example blood or bone marrow, or can be obtained from cell culture. After transfection the sample may be returned or readministered to a patient according to methods known to those practised in the art, for example, methods as described in Kasid *et al.*, Proc. Natl. Acad. Sci. USA (1990) **87**:473; Rosenberg *et al.* (1990) New Eng. J. Med. **323**: 570 ; Williams *et al.* (1994) Nature **310**: 476; Dick *et al.* (1985) Cell **42**:71; Keller *et al.* (1985) Nature **318**: 149 and Anderson *et al.* (1994) US Patent N. 5,399,346.

There are a number of viral vectors known to those skilled in the art which can be used to introduce the nucleic acid into mammalian cells, for example retroviruses, parvoviruses, coronaviruses, negative strand RNA viruses such as picornaviruses or alphaviruses and double stranded DNA viruses including adenoviruses, herpesviruses such as Herpes Simplex virus types 1 and 2, Epstein-Barr virus or cytomegalovirus and poxviruses such as vaccinia fowlpox or canarypox. Other viruses include, for example, Norwalk viruses, togaviruses, flaviviruses, reoviruses, papovaviruses, hepadnaviruses and hepatitis viruses.

A preferred method to introduce nucleic acid that encodes the desired protein into cells is through the use of engineered viral vectors. These vectors provide a means to introduce nucleic acids into cycling and quiescent cells and have been modified to reduce cytotoxicity and to improve genetic stability. The preparation and use of engineered Herpes simplex virus type 1 (D.M. Krisky, *et al.* (1997) Gene Therapy

4(10): 1120-1125), adenoviral (A. Amalfitanl, *et al.*(1998) Journal of Virology **72**(2):926-933), attenuated lentiviral (R. Zufferey, *et al.*, Nature Biotechnology (1997) **15**(9)871-875) and adenoviral/retroviral chimeric (M. Feng, et al, Nature Biotechnology (1997) **15**(9):866-870) vectors are known to the skilled artisan.

The protein may be administered using methods known in the art. For example, the mode of administration is preferably at the location of the target cells. The administration can be by injection. Other modes of administration (parenteral, mucosal, systemic, implant, intraperitoneal, etc.) are generally known in the art. The agents can, preferably, be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringer's solution and isotonic sodium chloride solution.

In yet another of its aspects the invention provides assay methods for identifying compounds that are able to enhance or inhibit the expression of the proteins of the invention. These assays can be conducted, for example, by transfecting a nucleic acid of the invention into host cells and then comparing the levels of mRNA transcript or the levels of protein expressed from said nucleic acids in the presence or absence of the compound.

Different methods, well known to those skilled in the art can be employed in order to measure transcription or expression levels.

Alternatively, it is possible to identify compounds that modulate transcription by using a reporter gene assay of the type well known in the art. In such an assay a reporter plasmid is constructed in which the

promoter of a gene, whose levels of transcription are to be monitored, is positioned upstream of a gene capable of expressing a reporter molecule. The reporter molecule is a molecule whose level of expression can be easily detected and may be either the transcript of the reporter gene or a protein with characteristics that allow it to be detected. For example, the molecule may be a fluorescent protein such as green fluorescent protein (GFP).

Compound assays may be conducted by introducing the reporter plasmid described above into an appropriate host cell and then measuring the amount of reporter molecule expressed in the presence or absence of the compound to be tested.

The invention also relates to compounds identified by the above mentioned methods.

Further embodiments of the present invention relate to methods of identifying the relevant gene or genes which involve the sub-cloning of YAC DNA as defined above into vectors such as BAC (bacterial artificial chromosome) or PAC (P1 or phage artificial chromosome) or cosmid vectors such as exon-trap cosmid vectors. The starting point for such methods is the construction of a contig map of the region of human chromosome 18q between polymorphic markers D18S60 and D18S61. To this end the present inventors have sequenced the end regions of the fragment of human DNA in each of the seven aforementioned YAC clones and these sequences are disclosed herein. Following subcloning of YAC DNA into other vectors as described above, probes comprising these end sequences or portions thereof, in particular those sequences shown in Figures 1 to 11 herein, together with any known

sequenced tagged site (STS) in this region, as described in the YAC clone contig shown herein, as can be used to detect overlaps between said subclones and a contig map can be constructed. Also the known sequences in the current YAC contig can be used for the generation of contig map subclones.

One route by which a gene or genes which is associated with a mood disorder or associated disorder can be identified is by use of the known technique of exon trapping.

This is an artificial RNA splicing assay, most often making use in current protocols of a specialized exon-trap cosmid vector. The vector contains an artificial minigene consisting of a segment of the SV40 genome containing an origin of replication and a powerful promoter sequence, two splicing-competent exons separated by an intron which contains a multiple cloning site and an SV40 polyadenylation site.

The YAC DNA is subcloned in the exon-trap vector and the recombinant DNA is transfected into a strain of mammalian cells. Transcription from the SV40 promoter results in an RNA transcript which normally splices to include the two exons of the minigene. If the cloned DNA itself contains a functional exon, it can be spliced to the exons present in the vector's minigene. Using reverse transcriptase a cDNA copy can be made and using specific PCR primers, splicing events involving exons of the insert DNA can be identified. Such a procedure can identify coding regions in the YAC DNA which can be compared to the equivalent regions of DNA from a person afflicted with a mood disorder or related disorder to identify the relevant gene.

Accordingly, in a further aspect the invention

comprises a method of identifying at least one human gene, including mutated variants and polymorphisms thereof, which is associated with a mood disorder or related disorder which comprises the steps of:

(a) transfecting mammalian cells with exon trap cosmid vectors prepared and mapped as described above;

(b) culturing said mammalian cells in an appropriate medium;

(c) isolating RNA transcripts expressed from the SV40 promoter;

(d) preparing cDNA from said RNA transcripts;

(e) identifying splicing events involving exons of the DNA subcloned into said exon trap cosmid vectors to elucidate positions of coding regions in said subcloned DNA;

(f) detecting differences between said coding regions and equivalent regions in the DNA of an individual afflicted with said mood disorder or related disorder; and

(g) identifying said gene or mutated or polymorphic variant thereof which is associated with said mood disorder or related disorders.

As an alternative to exon trapping the YAC DNA may be subcloned into BAC, PAC, cosmid or other vectors and a contig map constructed as described above. There are a variety of known methods available by which the position of relevant genes on the

subcloned DNA can be established as follows:

(a) cDNA selection or capture (also called direct selection and cDNA selection): this method involves the forming of genomic DNA/cDNA heteroduplexes by hybridizing a cloned DNA (e.g. an insert of a YAC DNA), to a complex mixture of cDNAs, such as the inserts of all cDNA clones from a specific (e.g. brain) cDNA library. Related sequences will hybridize and can be enriched in subsequent steps using biotin-streptavidine capturing and PCR (or related techniques);

(b) hybridization to mRNA/cDNA: a genomic clone (e.g. the insert of a specific cosmid) can be hybridized to a Northern blot of mRNA from a panel of culture cell lines or against appropriate (e.g. brain) cDNA libraries. A positive signal can indicate the presence of a gene within the cloned fragment;

(c) CpG island identification: CpG or HTF islands are short (about 1 kb) hypomethylated GC-rich (> 60%) sequences which are often found at the 5′ ends of genes. CpG islands often have restriction sites for several rare-cutter restriction enzymes. Clustering of rare-cutter restriction sites is indicative of a CpG island and therefore of a possible gene. CpG islands can be detected by hybridization of a DNA clone to Southern blots of genomic DNA digested with rare-cutting enzymes, or by island-rescue PCR (isolation of CpG islands from YACs by amplifying sequences between islands and neighbouring *Alu*-repeats);

(d) zoo-blotting: hybridizing a DNA clone (e.g.

the insert of a specific cosmid) at reduced stringency against a Southern blot of genomic DNA samples from a variety of animal species. Detection of hybridization signals can suggest conserved sequences, indicating a possible gene.

Accordingly, in a further aspect the invention comprises a method of identifying at least one human gene including mutated and polymorphic variants thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) subcloning the YAC DNA as described above into a cosmid, BAC, PAC or other vector;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other sequenced tagged site (STS) in this region as in the YAC clone contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to detect overlaps amongst the subclones and construct a map thereof;

(c) identifying the position of genes within the subcloned DNA by one or more of CpG island identification, zoo-blotting, hybridization of the subcloned DNA to a cDNA library or a Northern blot of mRNA from a panel of culture cell lines;

(d) detecting differences between said genes and equivalent region of the DNA of an individual afflicted with a mood disorder or related disorder; and

(e) identifying said gene which is associated

with said mood disorders or related disorders.

If the cloned YAC DNA is sequenced, computer analysis can be used to establish the presence of relevant genes. Techniques such as homology searching and exon prediction may be applied.

Once a candidate gene has been isolated in accordance with the methods of the invention more detailed comparisons may be made between the gene from a normal individual and one afflicted with a mood disorder such as a bipolar spectrum disorder. For example, there are two methods, described as "mutation testing", by which a mutation or polymorphism in a DNA sequence can be identified. In the first the DNA sample may be tested for the presence or absence of one specific mutation but this requires knowledge of what the mutation might be. In the second a sample of DNA is screened for any deviation from a control (normal) DNA. This latter method is more useful for identifying candidate genes where a mutation is not identified in advance.

In addition, the following techniques may be further applied to a gene identified by the above-described methods to identify differences between genes from normal or healthy individuals and those afflicted with a mood disorder or related disorder:

(a) Southern blotting techniques: a clone is hybridized to nylon membranes containing genomic DNA digested with different restriction enzymes of patients and healthy individuals. Large differences between patients and healthy individuals can be visualized using a radioactive labelling protocol;

(b) heteroduplex mobility in polyacrylamide gels: this technique is based on the fact that the mobility of heteroduplexes in non-denaturing polyacrylamide gels is less than the mobility of homoduplexes. It is most effective for fragments under 200 bp;

(c) single-strand conformational polymorphism analysis (SSCP or SSCA): single stranded DNA folds up to form complex structures that are stabilized by weak intramolecular bonds. The electrophoretic mobilities of these structures on non-denaturing polyacrylamide gels depends on their chain lengths and on their conformation;

(d) chemical cleavage of mismatches (CCM): a radiolabelled probe is hybridized to the test DNA, and mismatches detected by a series of chemical reactions that cleave one strand of the DNA at the site of the mismatch. This is a very sensitive method and can be applied to kilobase-length samples;

(e) enzymatic cleavage of mismatches: the assay is similar to CCM, but the cleavage is performed by certain bacteriophage enzymes.

(f) denaturing gradient gel electrophoresis: in this technique, DNA duplexes are forced to migrate through an electrophoretic gel in which there is a gradient of increasing amounts of a denaturant (chemical or temperature). Migration continues until the DNA duplexes reach a position on the gel wherein the strands melt and separate, after which the denatured DNA does not migrate much further. A single base pair difference between a normal and a mutant DNA duplex is sufficient to cause them to migrate to

different positions in the gel;

(g) direct DNA sequencing.

It will be appreciated that with respect to the methods described herein, in the step of detecting differences between coding regions from the YAC and the DNA of an individual afflicted with a mood disorder or related disorder, the said individual may be anybody with the disorder and not necessary a member of family MAD31.

In accordance with further aspects the present invention provides an isolated human gene and variants thereof associated with a mood disorder or related disorder and which is obtainable by any of the above described methods, an isolated human protein encoded by said gene and a cDNA encoding said protein.

In the experimental report which follows reference will be made to the following figures:

FIGURE 1 shows a sequence of nucleotides which is the left arm end-sequence of YAC 766_f_12;

FIGURE 2 shows a sequence of nucleotides which is a right arm end-sequence of YAC 766_f_12;

FIGURE 3 shows a sequence of nucleotides which is a left arm end-sequence of YAC 717_d_3;

FIGURE 4 shows a sequence of nucleotides which is a right arm end-sequence of YAC 717_d_3;

FIGURE 5 shows a sequence of nucleotides which is

a right arm end-sequence of YAC 731_c_7;

FIGURE 6 shows a sequence of nucleotides which is a left arm end-sequence of YAC 752_g_8;

FIGURE 7 shows a sequence of nucleotides which is a left arm end-sequence of YAC 942_c_3;

FIGURE 8 shows a sequence of nucleotides which is a right arm end-sequence of YAC 942_c_3;

FIGURE 9 shows a sequence of nucleotides which is a left arm end-sequence of YAC 961_h_9;

FIGURE 10 shows a sequence of nucleotides which is a right arm end-sequence of YAC 961_h_9;

FIGURE 11 shows a sequence of nucleotides which is a left arm end-sequence of YAC 907_e_1;

FIGURE 12 shows a pedigree of family MAD31;

FIGURE 13 shows the haplotype analysis for family MAD13. Affected individuals are represented by filled diamonds, open diamonds represent individuals who were asymptomatic at the last psychiatric evaluation. Dark gray bars represent markers for which it cannot be deduced if they are recombinant; and

FIGURE 14 shows the YAC contig map of the region of human chromosome 18 between the polymorphic markers D18560 and D18561. Black lines represent positive hits. YACs are not drawn to scale.

FIGURE 15 shows (a) a CAG repeat (in bold) and

surrounding nucleotide sequence isolated from YAC 961_h_9. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 16 shows (a) a CAG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 766_f_12. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 17 shows (a) a CAG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 766_f_12. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 18 shows (a) a CTG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 907_e_1. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

Experimental 1

(a) Family Data

Clinical diagnoses in MAD31, a Belgian family with a BPII proband were described in detail in De bruyn et al 1996. In that study only the 15 family members who

were informative for linkage analysis were selected for additional genotyping. The different clinical diagnoses in the family were as follows:
1 BPI, 2 BPII, 2UP, 4 Major depressive disorder (MDD), 1 SAm and 1 SAd.
The pedigree of the MAD31 family is shown in Figure 12.

(b) Genotyping of Family Members

All short tandem repeat (STR) genetic markers are di- or tetranucleotide repeat polymorphisms. Information concerning the genetic markers used in this study was obtained from several sources on the internet: Genome DataBase (GDB, http://gdbwww.gdb.org/), GenBank (http://www.ncbi.nlm.nih.gov/), Cooperative Human Linkage Center (CHLC, http://www.chlc.org/), Eccles Institute of Human Genetics (EIHG, http://www.genetics.utah.edu/) and Généthon (http://www.genethon.fr/). Standard PCR was performed in a 25 μl volume containing 100 ng genomic DNA, 200 mM of each dNTP, 1.25 mM $MgCl_2$, 30 pmol of each primer and 0.2 units Goldstar DNA polymerase (Eurogentec). One primer was end-labelled before PCR with [gamma-^{32}P]ATP and T4 polynucleotide kinase. After an initial denaturation step at 94°C for 2 min, 27 cycles were performed at 94°C for 1 min, at the appropriate annealing temperature for 1.5 min and extension at 72°C for 2 min. Finally, an additional elongation step was performed at 72°C for 5 min. PCR products were detected by electrophoresis on a 6% denaturing polyacrylamide gel and by exposure to an X-ray sensitive film. Successfully analysed STSs, STRs and ESTs covering the refined candidate region are fully described herein on pages 36 to 54.

(c) Lod score analysis.

Two-point lod scores were calculated for 3 different disease models using Fastlink 2.2. (Cottingham et al. 1993). For all models, a disease gene frequency of 1% and a phenocopy rate of 1/1000 was used. Model 1 included all patients and unaffected individuals with the latter individuals being assigned to a disease penetrance class depending on their age at examination. The 9 age-dependent penetrance classes as described by De bruyn et al (1996) were multiplied by a factor 0.7 corresponding to a reduction of the maximal penetrance of 99% to 70% for individuals older than 60 years (Ott 1991). Model 2 is similar to model 1, but patients were assigned a diagnostic stability score, calculated based on clinical data such as the number of episodes, the number of symptoms during the worst episode and history of treatment (Rice et al. 1987, De bruyn et al. 1996). Model 3 is as model 1 but includes only patients.

(d) Construction of the YAC contig - protocols

Growing of YACs and extraction of YAC DNA was done according to standard protocols (Silverman, 1995). For the construction of the YAC-contig spanning the chromosome 18q candidate region, the data of the physical map based on sequence tagged sites (STSs) (Hudson et al. 1995) was consulted on the Whitehead Institute (WI) Internet site (http://www-genome.wi.mit.edu/). CEPH mega-YACs were obtained from the YAC Screening Centre Leiden (YSCL, the Netherlands) and from CEPH (Paris, France). The YACs were analyzed for the presence of STSs and STRs, previously located between D18S51 and D18S61, by

touchdown PCR amplification. Information on the STSs/STRs was obtained from the WI, GDB, Généthon, CHLC and GenBank sites on the Internet. Thirty PCR cycles consisted of: denaturation at 94°C for 1 min, annealing (2 cycles for each temperature) starting from 65°C and decreasing to 51°C for 1.5 min and extension at 72°C for 2 min. This was followed by 10 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1.5 min and extension at 72°C for 2 min. A final extension step was performed for 10 min at 72°C. Amplified products were visualised by electrophoresis on a 1% TBE agarose gel and ethidium bromide staining.

(e) Ordering of the STR markers.

Twelve STR markers, previously located between D18S51 and D18S61, were tested for cosegregation with bipolar disease in family MAD31. The parental haplotypes were reconstructed from genotype information of the siblings in family MAD31 and minimalizing the number of possible recombinants. The result of this analysis is shown in Figure 13. The father was not informative for 3 markers, the mother was not informative for 5 markers. Haplotypes in family MAD31 suggested the following order for the STR markers analysed: cen-[S51-S68-S346]-[S55-S969-S1113-S483-S465]-[S876-S477]-S979-[S466-S817-S61]-tel. The order relative to each other of the markers between brackets could not be inferred from our haplotype data. The marker order in family MAD31 was compared with the marker order obtained using different mapping techniques and the results shown in Table 1 below.

Table 1. Comparison of the order of the markers within the 18q candidate region for bipolar disorder, among several maps.

Marker*	Genetic maps		Radiation hybrid map
	Généthon	Marshfield	(Giacalone et al. 1996)
D18S51		(-)3.4cM	(-)27.9 cR
D18S68	0 cM	0 cM	0 cR
D18S346		5.3 cM	52.2 cR
D18S55	0.1 cM	0 cM	72.5 cR
D18S969		0.6 cM	
D18S1113	0.7 cM		
D18S483	2.5 cM	3.2 cM	88 cR
D18S465	4.5 cM	5.3 cM	101.3 cR
D18S876			
D18S477	4.4 cM	5.3 cM	166.4 cR
D18S979		8.9 cM	
D18S466	7.6 cM	11.1 cM	212.4 cR
D18S61	8.4 cM	11.8 cM	249.5 cR
D18S817		5.3 cM	260.6 cR

* Order according to haplotyping results in family MAD31.

(-) Marker is located proximal of D18S68.

D18S68, common to all 3 maps, was taken as the map anchor point, and the genetic distance in cM or cR of the other markers relative to D18S68 are given. The marker order is in good agreement with the order of the markers on the recently published chromosome 18 radiation hybrid map (Giacalone et al. (1996) Genomics 37:9-18) and the WI YAC-contig map (http://www-genome.wi.mit.edu/). However, a few discrepancies with other maps were observed. The only discrepancy with the Généthon genetic map is the reversed order of D18S465 and D18S477. Two discrepancies were observed with the Marshfield map (http://www.marshmed.org/genetics/). The present inventors mapped D18S346 above D18S55 based on maternal haplotypes, but on the Marshfield maps D18S346 is located between D18S483 and D18S979. The inventors also placed D18S817 below D18S979, but on the Marshfield map this marker is located between D18S465 and D18S477. However, the location of D18S346 and D18S817 is in agreement with the chromosome 18 radiation hybrid map of Giacalone et al. (1996). One discrepancy was also observed with the WI radiation hybrid map (http://www-genome.wi.mit.edu/), in which D18S68 was located below D18S465. However, the inventors as well as other maps placed this marker above D18S55.

(f) Lod score analysis and refinement of the candidate region.

Lod score analysis gave positive results with all markers, confirming the previous observation that 18q21.33-q23 is implicated in BP disease, at least in family MAD31 (De bruyn et al. 1996). Summary statistics of the lod score analysis under all models

are given in table 2 below.

Table 2. Summary statistics of the two-point lod scores in MAD31.

Marker	Model 1			Model 2			Model 3		
	Z at θ=0.0	Zmax	θmax	Z at θ=0.0	Zmax	θmax	Z at θ=0.0	Zmax	θmax
D18S51	-0.19	0.73	0.1	0.94	0.94	0.01	0.08	0.54	0.1
D18S68	-0.19	0.73	0.1	0.94	0.94	0.01	0.07	0.55	0.1
D18S346	-0.19	0.73	0.1	0.94	0.94	0.01	0.07	0.55	0.1
D18969	1.40	1.40	0.0	1.27	1.27	0.0	1.20	1.20	0.0
D18S1113	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S876	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S477	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S979	-0.18	0.77	0.1	1.08	1.08	0.0	0.08	0.54	0.0
D18S817	-0.19	0.73	0.1	1.08	1.08	0.0	0.07	0.55	0.1
D18S61	-0.21	0.73	0.1	1.08	1.08	0.0	0.07	0.54	0.1

D18S55, D18S483, D18S465 and D18S466 were not informative.

The highest two-point lod score (+2.01 at θ=0.0) was obtained with markers D18S1113, D18S876 and D18S477 under model 1 in the absence of recombinants (table 2). In model 1, all individuals with a BP spectrum disorder are considered affected and fully contributing to the linkage analysis.
Before the fine mapping the candidate region was flanked by D18S51 and D18S61, which are separated by a genetic distance of 15.2 cM on the Marshfield map or 13.1 cM on the Généthon map. The informative recombinants with D18S51 and D18S61 were observed in 2 affected individuals (II.10 and II.11 in Fig. 13). However, since no other markers were tested within the candidate region it was not known whether these individuals actually shared a region identical-by-descent (IBD). The additional genetic mapping data now indicate that all affected individuals are sharing alleles at D18S969, D18S1113, D18S876 and D18S477 (Fig. 13, boxed haplotype). Also, alleles from markers D18S483 and D18S465 are probably IBD, but these markers were not informative in the affected parent I.1. Obligate recombinants were observed with the STR markers D18S68, D18S346, D18S979 and D18S817 (Table 2, fig. 13) Since discrepancies between different maps were observed for the locations of D18S346 and D18S817, the present inventors used D18S68 and D18S979 to redefine the candidate region for BP disease. The genetic distance between these 2 markers is 8.9 cM based on the Marshfield genetic map (http.//www.marshmed.org/genetics/).

(g) Construction of the YAC contig.

According to the WI integrated map 56 CEPH megaYACs are located in the initial candidate region

contained between D18S51 and D18S61 (Chumakov et al. (1995) Nature 377 Suppl., De bruyn et al. (1996)). From these YACs, those were selected that were located in the region between D18S60 and D18S61. D18S51 is not presented on the WI map, but is located close to D18S60 according to the Marshfield genetic map (http.//www.marshmed.org/genetics/). To limit the number of potential chimaeric YACs,YACs were eliminated that were also positive for non-chromosome 18 STSs. As such, 25 YACs were selected (see Figure 14), and placed in a contig based on the technique of YAC contig mapping, i.e. sequences from sequence tagged sites (STSs), simple tandem repeats (STRs) and expressed sequenced tags (ESTs), known to map between D18S60 and D18S61, were amplified by PCR on the DNA from the YAC clones. The STS, STR and EST sequences used, are described from page 36 to 54. Positive YAC clones were assembled in a YAC contig map (Figure 14).

Three gaps remained in the YAC contig, of which one, between D18S876 and GCT3G01, was located in the refined candidate region. To close the gap between D18S876 and GCT3G01, 14 YAC clones (Table 3, on page 62) were further analysed. End fragments from YAC clones 766_f_12 (SV11R), 752_g_8 (SV31L), 942_c_3 (SV10R) were obtained and sequenced (see pages 55-61). Primers from these three sequences were selected, and DNA of each of the 14 YAC clones was amplified by PCR. As indicated in Table 3, overlaps were obtained between 7 YAC clones on the centromeric side, and two YAC clones on the telomeric site (717_d_3 and 907_e_1).

The final YAC contig is shown in Figure 14. In the figure, only the YAC clones which rendered unambiguous hits with the chromosome 18 STSs, STRs and ESTs are shown. In a few cases, weak positive signals were also obtained with some of the YAC clones, which

likely represent false positive results. However, these signals did not influence the alignment of the YAC clones in the contig. Although, all YACs known to map in the region were tested as well as all available STSs/STRs, initially, the gap in the YAC contig was not closed. However, this was subsequently achieved by determining the end-sequences of the eight selected YACs (see below). The order of the markers provided by the YAC contig map is in complete agreement with the marker order provided by the WI map which integrates information from the genetic map, the radiation hybrid map and the STS YAC contig map (Hudson et al. 1995). Also, the YAC contig map confirms the order of the STR markers as suggested by the haplotype analysis in family MAD31. Moreover, the YAC contig map provides additional information on the relative order of the STR markers. For example, D18S55 is present in YAC 931_g_10 but not in 931_f_1 (Fig.14), separating D18S55 from its cluster [S55-S969-S1113-S483-S465] obtained by haplotype analysis in family MAD31. The centromeric location of D18S55 is defined by the STS/STR content of surrounding YACs (Fig. 14). If we combine the haplotype data and the YAC contig map the following order of STR markers is obtained: cen-[S51-S68-S346]-S55-[S969-S1113]-[S483-S465]-S876-S477-S979-S466-[S817-S61]-tel.

Out of the 25 YAC clones spanning the whole contig, seven YAC clones were selected in order to identify the minimal tiling path (Table 4). These 7 YAC clones cover the whole refined chromosome 18 region. Furthermore, YAC clones should preferably be non-chimeric, i.e. they should only contain fragments from human chromosome 18. In order to examine for the presence of chimerism, both ends of these YACs were subcloned and sequenced (pages 55 to 61). For each of

the sequences, primers were obtained, and DNA from a monochromosomal mapping panel was amplified by PCR using these primers. As indicated on pages 55 to 61, some of the YAC clones contained fragments from other chromosomes, apart from human chromosome 18.

Three YAC clones were then selected comprising the minimum tiling path (Table 5). These three YAC clones were stable as determined by pulsed field gel electrophoresis and their seizes correspond well to the published sizes. These YAC clones were transferred to other host yeast strains for restriction mapping, and are the subject to further subcloning.

Description of the succesfully analysed STSs, STRs and ESTs covering the refined candidate region.

Explanations:

- STS: Sequence Tagged Site
- STR: Simple Tandem Repeat
- EST: Expressed Sequence Tag

These markers are ordered from the centromere to the telomere. Only the markers that were effectively tested and that worked on the YACs are given.

List:

1. D18S60:

Database ID: AFM178XE3 (Also known as 178xe3, Z16781, D18S60)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = CCTGGCTCACCTGGCA
Right = TTGTAGCATCGTTGTAATGTTCC
Product Length = 157
Review complete sequence:
AGCTATCCTGGCTCACCTGGCAAAAATACAGTGTATACACACACACACAC
ACACACACACACACACACAGAGTGTNTTANTNATTCCAGCAAATAATATTA
CATATAAAAGATCTAATTGGTTCATCATGTAAATTTAGTAGGAACATTACA
ACGATGCTACAAGANTTTATCCAAAACTGAGATTTCCTTAGAATATCTGTT
AAAAGTAATTTTATTCAGTTAATAGAAATTCTATTGAAAACATCAAACTTAT
AAAGCT

Genbank ID: Z16781
Description: H. sapiens (D18S60) DNA segment containing (CA) repeat; clone
Search for GDB entry

2. WI-9222:

Database ID: UTR-03540 (Also known as G06101, D18S1033, 9222, X63657)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:

Left = GATCCCATAAAGCTACGAGGG
Right = GAGTCTAAAGACAAGAAAGCATTGC
Product Length = 99
Review complete sequence:

TCTTCTTACCCCTTGGAAGAAGACTGTTTCCAAATAATTTGAACAGCTTG
CTGCTAAATGGGACCCAATTTTTGGCCTATAGACACTTATGTATTGTTTTC
GAATACGTCAGATTGGACCAGTGCTCTTCAGGAATGTGGCTGCAAGCAA
GGGGCTAGAAGTTCACCTCCTGACAGTATTATTAATACTATGCAAATATG
GAATAGGAGACCATTTGATTTTCTAGGCTTTGTGGTAGAGAGGTGAAGG
TATGAGAATTAATAGCGTGTGAACAAAGTAAAGAACAGGATTCCAGAATG
ATCATTAAATTTGTTTCTATTTATTCTTTTTTGCCCCCCTAGAGATTAAGTC
CAGAAATGTACTTTCTGGCACATAAAGAAATCTTGAGGACTTTGTTTAAAC
CTTCCATAAAAAAAACAATTTTCGGTTTCTCGGGTNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT
TCTTTCTTTGTGTATTTTATTCAAGATGAGTTGGACCCATTGCCAGTGAGT
CTGAATGTCACTGACAGCCCTGTGTTGTGCTCAGGACTCACTCTGCTGC
TGGTGGAAACTCATGGCTTCTCTCTCTCTTTGATCCCATAAAGCTACGAG
GGGGACGGGAGAGGGCAGTGCAATGGGAAGTAAAGAGATATTTTCCAG
TAGGAAAAGCAATGCTTTCTTGTCTTTAGACTCAAATGCTTAGGGAACGT
TTCATTTCTCATTCATGGGGAAAGGCAGCCTCCTTAAATGTTTTCTGAAG
AGCGGTAAAATCTAGAAGCTTAAGAATTTACAGTTCCTTCAATAACCATGA
TGACCTGAAGTTCACCTATCCCATTTTAGCATCTACTTGTTTTTCCCATCT
CTTCCTTTCCAATTTTGCTTATACTGCTGTAATATTTTTGTNNNNNNNNNN
NNNNNNNNNNNNNNGACCAGCTAAAATTTTCGACTTGACTTTTTAACTT
AACTCATGAATTAATTAAAGCAAATGAAAAAATTAAAAAGTGTGACTTTTT
CTCGGAGCATATATGTAGCTTTTAGGAAAGGCTGATGATGGTATAAAGTT
TGCTCATTAAGAAAAAAAGACAAGGCTGATTTTGAAGAGAGTTGCTTTTG
AAATAAAATGATCA

Genbank ID: X63657
Description: H.sapiens fvt1 mRNA
Search for GDB entry

3. WI-7336:

Database ID: UTR-04664 (Also known as PI5, G00-679-135, G06527, 7336, U04313)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = AGACATTCTCGCTTCCCTGA
Right = AATTTTGACCCCTTATGGGC
Product Length = 332
Review complete sequence:

TAAGTGGCATAGCCCATGTTAAGTCCTCCCTGACTTTTCTGTGGATGCCG
ATTTCTGTAAACTCTGCATCCAGAGATTCATTTTCTAGATACAATAAATTG
CTAATGTTGCTGGATCAGGAAGCCGCCAGTACTTGTCATATGTAGCCTTC

ACACAGATAGACCNNNNNNNNNNCCAATTCTATCTTTTGTTTCCTTTTTT
CCCATAAGACAATGACATACGCTTTTAATGAAAAGGAATCACGTTAGAGG
AAAAATATTTATTCATTATTTGTCAAATTGTCCGGGGTAGTTGGCAGAAAT
ACAGTCTTCCACAAAGAAAATTCCTATAAGGAAGATTTGGAAGCTCTTCT
TCCCAGCACTATGCTTTCCTTCTTTGGGATAGAGAATGTTCCAGACATTC
TCGCTTCCCTGAAAGACTGAAGAAAGTGTAGTGCATGGGACCCACGAAA
CTGCCCTGGCTCCAGTGAAACTTGGGCACATGCTCAGGCTACTATAGGT
CCAGAAGTCCTTATGTTAAGCCCTGGCAGGCAGGTGTTTATTAAAATTCT
GAATTTTGGGGATTTTCAAAAGATAATATTTTACATACACTGTATGTTATA
GAACTTCATGGATCAGATCTGGGGCAGCAACCTATAAATCAACACCTTAA
TATGCTGCAACAAAATGTAGAATATTCAGACAAAATGGATACATAAAGACT
AAGTAGCCCATAAGGGGTCAAAATTTGCTGCCAAATGCGTATGCCACCA
ACTTACAAAAACACTTCGTTCGCAGAGCTTTTCAGATTGTGGAATGTTGG
ATAAGGAATTATAGACCTCTAGTAGCTGAAATGCAAGACCCCAAGAGGAA
GTTCAGATCTTAATATAAATTCACTTTCATTTTTGATAGCTGTCCCATCTG
GTCATGTGGTTGGCACTAGACTGGTGGCAGGGGCTTCTAGCTGACTCG
CACAGGGATTCTCACAATAGCCGATATCAGAATTTGTGTTGAAGGAACTT
GTCTCTTCATCTAATATGATAGCGGGAAAAGGAGAGGAAACTACTGCCTT
TAGAAAATATAAGTAAAGTGATTAAAGTGCTCACGTTACCTTGACACATAG
TTTTTCAGTCTATGGGTTTAGTTACTTTAGATGGCAAGCATGTAACTTATA
TTAATAGTAATTTGTAAAGTTGGGTGGATAAGCTATCCCTGTTGCCGGTT
CATGGATTACTTCTCTATAAAAAATATATATTTACCAAAAAATTTTGTGACA
TTCCTTCTCCCATCTCTTCCTTGACATGCATTGTAAATAGGTTCTTCTTGT
TCTGAGATTCAATATTGAATTTCTCCTATGCTATTGACAATAAAATATTATT
GAACTACC

Genbank ID: G06527
Description: WICGR: Random genome wide STSs

4. WI-8145:

Database ID: EST102441 (Also known as D18S1234, G00-677-827, G06845, 8145, T49159)
Source: WICGR: STSs derived from dbEST sequences
Chromosome: Chr18

Primers:
Left = GAAATGCACATAACATATATTTGCC
Right = TGCTCACTGCCTATTTAATGTAGC
Product Length = 184
Review complete sequence:

GTTGTTTGGANGCAGGTTTATTTATTATATACTTGCAATTGAATATAAGAT
ACAGACATATATATGTGTTATGTATTTCTAGAAATGCACATAACATATATTT
GCCTATTGTTTAATGTTTTTTCCAGANATTTATTACAGAAGGGCATGGAG
GGATACCTACTTATTCTTCATTATGAGAACAATTAAAGGCATTTATTAGAT
AGGAAATTAACAGANCATCTGCTTCTATAACTTTATTAGCTACATTAAATA
GGCAGTGAGCANTAATTTAAAANCTCACCATTATATAAANTANTAAATACC
AAAGTAAAAG

_____: left and right primer

PCR Conditions

Genbank ID: T49159
Description: yb09e07.s1 Homo sapiens cDNA clone 70692 3' similar to gb:J02685
UniGene Cluster Description: Human mRNA for Arg-Serpin (plasminogen activator-inhibitor 2, PAI-2) Search for GDB entry

5. WI-7061:

Database ID: UTR-02902 (Also known as PAI2, G00-678-979, G06377, 7061, M18082)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = TGCTCTTCTGAACAACTTCTGC
Right = ATAGAAGGGCATGGAGGGAT
Product Length = 338
Review complete sequence:

AACTAAGCGTGCTGCTTCTGCAAAAGATTTTTGTAGATGAGCTGTGTGCC
TCAGAATTGCTATTTCAAATTGCCAAAAATTTAGAGATGTTTTCTACATAT
TTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAATAAAAACACAGAAAT
AATTAGACAATTGTCTATTATAACATGACAACCCTATTAATCATTTGGTCT
TCTAAAATGGGATCATGCCCATTTAGATTTTCCTTACTATCAGTTTATTTT
TATAACATTAACTTTTACTTTGTTATTTATTATTTTATATAATGGTGAGTTTT
AAATTATTGCTCACTGCCTATTTAATGTAGCTAATAAAGTTATAGAAGCAG
ATGATCTGTTAATTTCCTATCTAATAAATGCCTTTAATTGTTCTCATAATGA
AGAATAAGTAGGTATCCCTCCATGCCCTTCTATAATAAATATCTGGAAAAA
ACATTAAACAATAGGCAAATATATGTTATGTGCATTTCTAGAAATACATAA
CACATATATATGTCTGTATCTTATATTCAATTGCAAGTATATAATAAATAAA
CCTGCTTCCAAACAACNNNNNNNNNNNNNNGGAATTC

PCR Conditions

Genbank ID: G06377
Description: WICGR Random genome wide STSs

6. D18S68:

Database ID: AFM243YB9 (Also known as 248yb9, Z17122, D18S68)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATGGGAGACGTAATACACCC
Right = ATGCTGCTGGTCTGAGG
Product Length = 285
Review complete sequence:

AAAGAGTTGGGGTTGTGAATTCCCACACCAGTCAACTATTGGCTATGGG
CTTACCATGGGAGACGTAATACACCCGGNACTTCCAACTCACATACCAG
AGACATGGCTCTAGCACCCAATGGAAATATGCTGAATGTTGCAGGTGCA
AGACAGCAACAAAGCAGACAGAGGCACATAGACAAGGCACCAACAGTGT
CCACTATACCCTGACAGTGTGGAAAGTTGTAGATAGGATGAAGAGAAAG
AATACACACACACACACACACACACACACACACACACACACACACACACA
CGGTAGANACTTACTACNCAAAGTGTGANCCTCAGACCAGCAGCATCTG
GCNAAATGGTGATCTATCACCTTCCAG

Genbank ID: Z17122
Description: H. sapiens (D18S68) DNA segment containing (CA) repeat; clone

7. WI-3170:

Database ID: MR3726 (Also known as D18S1037, G04207, HALd22f2, 3170)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TGTGCTACTGATTAAGGTAAAGGC
Right = TGCTTCTTCAATTTGTAGAGTTGG
Product Length = 156
Review complete sequence

CTGAGACAAGGCAGGCAAACAACCTCTAAAAATCTACAATTGGTGATTGG
TGTGCTACTGATTAAGGTAAAGGCACAGAATTATACATCCAGGTTNCTAT
TACTTATGGCAGACTCAGGACCCAGGTTNAGAGACCACTGGCCTTAAGA
AAAAAAATGGGGTTCCTGATTTCTGGATAATAATCCAACTCTACAAATTGA
AGAAGCAACATACCCTCTTTGTTA

Genbank ID: G04207
Description: WICGR: Random genome wide STSs

8. WI-5654:

Database ID: MR10908 (Also known as D18S1259, G00-678-695, G05278, 5654)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = CTTAATGAAAACAATGCCAGAGC
Right = TGCAAAATGTGGAATAATCTGG
Product Length = 149
Review complete sequence:

CTACAAAATGCATGTGGCTTTGGCTTTGAAATAGTACACCCTATCAAAGA
CTAAATTTTCTTAATGAAAACAATGCCAGAGCTTTTTTCATGATATTTTGTT
TTTAGAGATGGGGAACAATCTGGACGTTGTTTCCTTATCTGGGTGGTAAT
CGAGGCTTAGCAATTTCCCACAGCGTTACACAAATCCAGATTATTCCACA
TTTTGCAAATA

Genbank ID: G05278
Description: WICGR: Random genome wide STSs

9. D18S55:

Database ID: AFM122XC1 (Also known as 122xc1, Z16621, D18S55, GC378-D18S55)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18
Primers:
Left = GGGAAGTCAAATGCAAAATC
Right = AGCTTCTGAGTAATCTTATGCTGTG
Product Length = 143
Review complete sequence:
AGCTGAACATGCCTTTTCATGGAGCAGTTTCNAAATACACTTTTGGTACA
ATCTGCAGGTGGATATTTGGAGCTCAGGAGTTTGAGACCAGCCTGGGCA
ACATGGTGAAATCCCGTCTCTACTAAAATACAAAAAATTAGCCAGGTGTG
GCGGCATGTGCCTGTAGNCCCAGGATGGATTGAGTGGGTGAGATATGG
AATAAGTGGTGGGAAGTCAAATGCAAAATCAATTCAGTTTGTCAATATTG
ATTCTCTATTCTAGCCTGGCGTGGTTTTTCCTCGTCACACACACACACAC
ACACACACACACACACACACACACACACACACACACACAGCATAAGATTACTCAGA
AGCT

Genbank ID: Z16621
Description: H. sapiens (D18S55) DNA segment containing (CA) repeat; clone

10. D18S969:

Database ID: GATA-P18099 (Also known as G08003, CHLC.GATA69F01, CHLC.GATA69F01.P18099)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AACAAGTGTGTATGGGGGTG

Right = CATATTCACCCAGTTTGTTGC
Product Length = 365
Review complete sequence:

CAGGGAAATGCAAATCAAAACCACAATGAGTTATCTCCTCATACCTTTAAT
GATGGCTAATATTAAACAAGAGATAACAAGTGTGTATGGGGGTGTGGAG
AAAAGAGAATGTNCGAACACTCTTGGTTGAAATATAAGTTGGTAGANCCA
TTATGCAAAACAGTATGAATCTTTATCAGTATAANATTAGGACCTNGCATA
TGATCNCAGCAATCNCCACNTCTGNGNGATCNCACNCNCTATCTCTCTAT
ATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT
ATCTGTCTGTCTATCATCTATCTATCTTCTATCTATCTATCTATCTTTCTAT
CTATCTATCTGTCTATCTATNCCGGAATATTTTTCAGCCATNNAAATAAGG
AAGTCCTGCTATTTGCAACAAACTGGGTGAATATGGAGAACGTTATGCTA
AATGCAATATGCTAAAGACAGACACAGAAAGACAAGTATGACCTCACTTA
TATGTGGAAACTGAAAAAGCCATACTCATTACAGCAAAGAGTAGAATGTT
GGTTACCAGGGGCAAAGAGGGTAGAAATGAGGGGAGTGAGAAAATGTC
AATCAAAGTGTAAGAATGTTATAACATAAATAAATTCATAGAG

Genbank ID: G08003
Description: human STS CHLC.GATA69F01.P18099 clone GATA69F01.

11. D18S1113:

Database ID: AFM200VG9 (Also known as D18S1113, 200vg9, w2403)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = GTTGACTCAAGTCCAAACCTG
Right = CAAAGACATTGTAGACGTTCTCTG
Product Length = 207
Review complete sequence:
AGCTGCATATAAAACTATTCCATTTCACATTTTTGAAGACATTTGTAGCCA
TGATACTTTGCTGTTGTCTGTGGGCCACCTCTTTTTGAAGTGTGTAGTTA
ACTGTGCTCCTGTAATCTGTTGTCTGTTGACTCAAGTCCAAACCTGTTCT
GCGTGGCATGTTTCTNCAACTTGATGTGATGCTATTTATCACTTTCTTTGA
AGTTAAGTCTCTATGTCTTTGTATTCTTTCTGTGTACCCAGGGATATGTTT
GTGCATGCACACGCATAAACACACACACACACACACACACACACAGAGA
CAGAGACAGAGAACGTCTACAATGTCTTTGTGAG

12. D18S868:

Database ID: GATA-D18S868 (Also known as G09150, CHLC.GATA3E12, CHLC.GATA3E12.496. CHLC.496, D18S868)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AGCCAATACCTTGTAGTAAATATCC
Right = GATTCTCCAGACAAATAATCCC
Product Length = 189
Review complete sequence:
GAGTGAGCCAATACCTTGTAGTAAATATCCATCTATCTTTGATGTATCTAT
GTATCTATCTTTGTATCTATATGTCTATGTATCTATGTATGTATGTATCTAT
CTATCATCTATCTATCTATCATCTATCTATCTATCTATCTATCTATCTATCT
ATCTATCTATATCCNTTTGGGATTATTTGTCTGGAGAATCCTGATTAACAT
AGTCTGCTAACTTTTATCTGTATCTCCTATGTGTATGCTTCTCCTTCTTCC
TGTCTCTCTCTCTTCTTTGTCCTCATTTAANCTCCTTTCCTGGGNATATTG
GNAATTTTGATTGGANTCTGGACANTGTAGGAGTAAAAATTT

Genbank ID: G09150
Description: human STS CHLC.GATA3E12.P6553 clone GATA3E12.

13. WI-9959:

Database ID: MR12816 (Also known as D18S1251, G00-678-524, G05488, 9959)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TGCCAACAGCAGTCAAGC
Right = AGCACCTGCAGCAGTAATAGC
Product Length = 110
Review complete sequence:
ctgttttatttgaaaaaaaaaatctgtctccaagaagaaaagttcattctACCTGTTGCCAACAGC
AGTCAAGCGGACATGTTTAAAATTTTTTAAAAAAGTATTTTTTTTTCCAACT
GGNGTTTAATAGCCTCATTTTGGCTTTTGCTATTACTGCTGCAGGTGCTT
TNATTTTTTTCCTCTGCATTATAATTAC

Genbank ID: G05488
Description: WICGR: Random genome wide STSs
Search for GDB entry

14. D18S537:

Database ID: CHLC.GATA2E06.13 (Also known as CHLC.13, GATA2E06, D18S537, GATA-D18S537)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCCATCTATCTTTGATGTATCTATG
Right = AGTTAGCAGACTATGTTAATCAGGA
Product Length = 191

Review complete sequence:
AAAGCTGAGTGAGCCAATACCTTGTAGTAAATATCCATCTATCTTTGATGT
ATCTATGTATCTATCTTTGTATCTATATGTCTATGTATCTATGTATGTATGT
ATCTATCTATCATCTATCTATCTATCATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATATCCNTTNGGTATTATTNGTCTGGNGAATCCTGAT
TAACATAGTCTGCTAACTTNTATCTGTATCTNCTATGTGTATGCTTCTNCT
TCTTCCTGTCTCTCTCTCTGCTTTGTCCTCAATTNAAATCTCC

Genbank ID: G07990
Description: human STS CHLC.GATA2E06.P6006 clone GATA2E06.
Search for GDB entry

15. D18S483:

Database ID: AFM324WC9 (Also known as 324wc9, Z24399, D18S483)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = TTCTGCACAATTTCAATAGATTC
Right = GAACTGAGCAAACGAGTATGA
Product Length = 214
Review complete sequence:
AGCTCTGCTGGAAGAGCAGGGCTGTTTTCTGCACAATTTCAATAGATTCC
CCTACCCTGGGTTTTTCAGTAGATAGATAGATAGATGATAGATAGGTAGA
TAGATAGATAGATAGATAGATAGATAGATAGATAGATGATAGATAGATTTT
ATATATAGTATATAAAATCTACACACACACACACACACACACACACACATA
TTTGCCTTTCCTTGACTATCATACTCGTTTGCTCAGTTCTTTTTTTTTTAA
ATTTTTGTTTGTAAATCCAAAATGCTT

Genbank ID: Z24399
Description: H. sapiens (D18S483) DNA segment containing (CA) repeat; clone
Search for GDB entry

16. D18S465:

Database ID: AFM260YH1 (Also known as 260yh1, Z23850, D18S465)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATATTCCCCTATGGAAGTACAG
Right = AAAGTTAATTTTCAGGCACTCT
Product Length = 232
Review complete sequence:
AGCTCTGTCCCTCTAGAGAACGCTGACTAATATATTCCCCTATGGAAGTA
CAGATGGTTTTTTNTAAAATAAATTTATCTGATTGTGATGAGATAATCATCA

TTTTTATGTTCAGTGTTTTTCTAAATTTTTTATTGTTATTGTTTTTATACTCT
AAATGGTTTTTAAATATGCACATATGTGCATATTTTACACACACACACACA
CACACACACACTCTCTTTATTTAGAAGCATTATAGATAGAGTGCCTGAAAA
TTAACTTTTAACCNAAGAAAAGACAATAAGGAACAATAGGGAAGTTATCC
TTTGCTAAGGGTATGGAAAATATTCACATATTATTTATAACANGTTAAACC
AAGTCATGCTTGANTATAATAGCT

Genbank ID: Z23850
Description: H. sapiens (D18S465) DNA segment containing (CA) repeat; clone
Search for GDB entry

17. D18S968:

Database ID: GATA-P34272 (Also known as G10262, CHLC.GATA117C05, CHLC.GATA117C05.P34272)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = GAAATTAACCAGACACTCCTAACC
Right = CTTAGAATTGCCTTTGCTGC
Product Length = 147
Review complete sequence:
GAATAAAAATATGAGGTATTAGAAATTTACAGATAGGAAGAAATTAACCAG
ACACTCCTAACCACCGATNAGTTTAAAGAGGAGATAGATAGATAGATGAT
AGATAGATAGATAGATAGATAGATACCACTGAAAATGCAANCACAAATTA
GCAGATTATATGTGATGCAGCAAAGGCAATTCTAAGTAGATTCTAACTGC
TACATTGATAGCAGTACCCACTGACATTACCGGAAAGGATGGTATCCATA
ACCACCTACCTATATACCTCCGCAGCTGGANATTAGGNTTAAGCTTCTTN
GGGCNCCTGGCGGCCCCNNTTGTGGTCCCCGGTNGGNCCCCGNTTNN
GNNTNGCTNNGNTTNCNTTGGNGNCCCCCNNTNGGTTTNNGGNNNNNT
NNNNNTNGNNNNNTTNCCCNNNNNNNNTNTNNNNCNNNNNNNNTNNN
NNNNNNNNNNGGNNNNNGGGN

Genbank ID: G10262
Description: human STS CHLC.GATA117C05.P34272 clone GATA117C05.

18. GATA-P6051:

Database ID: GATA-P6051 (Also known as CHLC.GATA3E08, CHLC.GATA3E08.P6051)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = GCAACAACCCTAATGAGTATACG

Right = GAGTCTCACCAGGGCTTACA
Product Length = 149
Review complete sequence:
AAAGCTGTCTCCTTTTGTAAAGTGTGCTCAGAGGAATCTTTTTCAGTAAAT
AAAGTCTGCACCCAGACATCTCACTTTGTATACCACGGAGAATTTACCAT
GACTCTTCTCAGTGATAAACGTCAATATAGAATAATCAGGAGAAAAAGAG
AAATCCAGTAAAGAAATAAGTCTGTAGAAAGCAACAACCCTAATGAGTAT
ACGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATG
NATCTATCTATCTAACATTATATAAAATATATATTTCCTCCTGTATTGGGG
CCCTGTGTGTAAGCCCTGGTGAGACTCAAAAATTTGANTATTCCTNTTTN
T

Genbank ID: G09104
Description: human STS CHLC.GATA3E08.P6051 clone GATA3E08.

19. D18S875:

Database ID: GATA-D18S875 (Also known as G08001, CHLC.GATA52H04, D18S875)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCCTCTCATCTCGGATATGG
Right = AAGGCTTTCAGACTTACACTGG
Product Length = 394
Review complete sequence:
TTATTTATTCACTCATTCAATAAATATTTATGAATTTCCTTTAATGGCNANG
AAAGTATGTTTGGTACTGAATATGGTGAGCAAGATTTTCCTCTCATCTCG
GATATGGAAAGATCTTGGAAATCATTATACNTCATACTTACAATANGAAAG
AAGCTGAGCAATTTGAAAATCAACAATTTCTTTTGTACNTGTCAGAAAAGT
GAAGATATATTAATCAGGGTTCTTCAGAGAAACATAACCAATAGGNCACA
GNTCTATATGNCCNCNTTTATCTATCTATCTATCTATCTATCNCTATCTAT
CNCANACCNGGNGAANTNATNTTTGNGAGATTNATGCAAGNCTGAGAAA
NACCNAAGAANCTGCTCCCTGTNAAACTNGAGATNCAAGAANCTGAANA
GTATAGNTCCAGTCCNAAGTCTANAGACCTTAGAATTAGGAAAACTGATA
CTATAAATACCAGTGTAAGTCTGAAAGCCTTAAANACCANATAGTGCCAT
TGAAAGGGCAGAAGACTGATGTCCCAGTTCAAGCAGGCAAAGTTAGAGA
AGCCTTATTTTCTGCAACATTGTTCTATTCAGACCCTTNANANGATTGACN
ATGTCCACCCA

Genbank ID: G08001
Description: human STS CHLC.GATA52H04.P16177 clone GATA52H04.
Search for GDB entry

20. WI-2620:

Database ID: MR1436 (Also known as G03602, D18S890, HHAa12h3, 2620)
Source: WICGR: Random genome wide STSs

Chromosome: Chr18

Primers:
Left = TCTCCAAGCTATTGATTGGATAA
Right = TTAAGAGCCAATTTATATAAAAGCAGC
Product Length = 177
Review complete sequence:
CCCCTTTTGCCAACGCCATGCTTCACGTAGGGAGCCTGACATGCAGAAA
ACTCTCCAAGCTATTGATTGGATAAAGAGCCAGAGCTGACTGAATTCCAT
TCTTCTTGAGCCTCTCATTCTGTGTTTCTCGAATTTTTACCAAAGCATCTT
GACACACAAATATCTGACTCAAGGAAAAGGAAAAACAACTGCTTTTTCTC
CAGCTGCTTTTATATAAATTGGCTCTTAAACTTTCTAAGTTTATTATGGAT
A

Genbank ID: G03602
Description: WICGR: Random genome wide STSs
Search for GDB entry

21. WI-4211:

Database ID: MR6638 (Also known as G03617, D18S980, 4211)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = ATGCTTCAGGATGACGTAATACA
Right = AAATTCTCGCTGATTGGAGG
Product Length = 113
Review complete sequence:
CTAGTACCATAATCCCTTTTGGAATAAACCATCCCACCTTTAGTCAGANC
AGATGCTTCAGGATGACGTAATACATAATAAGCCTACTCAGTTCTACTCT
GGCTTTGTATGTCTTCAAAGTGATATTTTTTTAAGTATTACTTGTCCCTCC
AATCAGCGAGAATTT

Genbank ID: G03617
Description: WICGR: Random genome wide STSs
Search for GDB entry

22. D18S876:

Database ID: GATA-D18S876 (Also known as G09963, CHLC.GATA61E10, D18S876)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCAAACTTATAACTGCAGAGAACG
Right = ATGGTAAACCCTCCCCATTA

Product Length = 171
Review complete sequence:
AAGACTGCAATTACATTTGCATCAAACTTATAACTGCAGAGAACGTTGCC
CACTATTTTATACCACACAACAGTATTCTTAGCCAGATTACATCTATCTAT
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCATCTATCTAGC
TAGCTATCTATCTATAGAACTAATGGGGAGGGTTTACCATGTTTGGGTGA
ACCCAAACATTTTATGGNCAAGGGNTTGGAAAATTACCCTTATCTACAAC
TNTTNAACTTGTTTTGGTAGGNGTGNTAATTCCNTGGGNTTGGAANAACT
TTTGNAATTTCCTCNTTGTTTNTNATTNNNNATTNNTNNNCATTATTNTGG
GGTNTTCNGGGTGGAGGGCTNANTTTGGCCNCCCGGGTCCNNGGNGC
NAGTNGGNNNGGNTNNTNGGGTTTNCTTGGGAANCNTNCCNCCTNCNG
GGGNTTCANGGGNTTTTTNTTTNNTTG

Genbank ID: G09963
Description: human STS CHLC.GATA61E10.P17745 clone GATA61E10.
Search for GDB entry

23. GCT3G01:

Database ID: GCT-P10825 (Also known as G09484, CHLC.GCT3G01, CHLC.GCT3G01.P10825)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = CTTTGCAATCTTAGTTAATTGGC
Right = GAACTATGATATGGAGTAACAGCG
Product Length = 128
Review complete sequence:
AGATGTTTAACTTTGCAATCTTAGTTAATTGGCAGAAATGAAATTTAGTTT
CCACAACTTTTATTCGATATTAAAACACCACCACCATCAGCAGCAGCAGC
AGCAGCAGCAGCATCGCTGTTACTCCATATCATAGTTCAGAGCATTTAAA
GNGGTCAAAATATACAACTAGGCTGACACCNGNATAAGGTTTAATTTTAA
ACCNGNGGTCTNCCCTCTAAGGNGGNTTTTTTTTTCTTGNCNTGGCTTCT
TTTTCCNTTTGCTTTTGTAAAATATCAAGGNATTTTTGGGTTNTTCNTGGN
ANTTNNCNNANTNNTNNTTNNNCNCNCCCCCCNTTTGNGGCGGGGGTC
CCCNNNTTGCCCCGGGGTTGNGTGCAGTAGGGGGGTCNCGGGTNNNG
NAAGTTTNGGGGCCCT

Genbank ID: G09484
Description: human STS CHLC.GCT3G01.P10825 clone GCT3G01.

24. WI-528:

Database ID: MH232 (Also known as G03589, 528, D18S828)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TTCTGCCTTTCCTGACTGTC
Right = TGTTTCCCATGTCTTGATGA
Product Length = 211
Review complete sequence:
CTACTAAGCAAATTCTGCTCAGCCTTCTGCCTTTCCTGACTGTCTTGTTG
GCCCTTCCCACTTTAAGGATGCCTGTTTAAGTAGCCACCTCTAATTAGGA
ATCTTCCCTTGTTCTTTCTCAGGAGGCTTAGACACTGTCAGTTTCCTGAA
GACAGAAAATAAGCCTGCATTATCCTAGTAGTGGATTCAAAACTAATTGT
GTCCTGAGTCTTTCAATCATCAAGACATGGGAAACACTCAACAG

Genbank ID: G03589
Description: WICGR: Random genome wide STSs
Search for GDB entry

25. WI-1783:

Database ID: MR432 (Also known as G03587, _shu_31.Seq, 1783, D18S824)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = CCAGTAATTAGACATTGACAGGTTC
Right = TTTTACTAGACAGGCTTGATAAACAA
Product Length = 305
Review complete sequence:
CCAGTAATTAGACATTGACAGGTTCCATACTAGTAATGTAGGGAATAGGG
CTGCTGCTTTTTGGGTTTCCTTGAGTATACTTTGTGCTGCATAAATATGG
CAATGGATAGTAAATAATTTGTATGCAGACCTTTAGTGTCGATTAACCTGT
GAATAAGGGAACAACAATCAAGGACAAAAATCAAAAGACTAATTCTCTAT
ACATTTTGAGCTTTTGTAAAAAAGTAAGATTAGCTGAATATATCTGAAAAA
TTTCTAATCTCCTTTACAATTTTTTAAATTGTTTATCAAGCCTGTCTAGTAA
AAATAATTCAGTTTCGGAATGTGG

Genbank ID: G03587
Description: WICGR: Random genome wide STSs
Search for GDB entry

26. D18S477:

Database ID: AFM301XF5 (Also known as 301xf5, Z24212, D18S477)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = GGACATCCTTGATTTGCTCATAA
Right = GATTGACTGAAAACAGGCACAT

Product Length = 243
Review complete sequence:
GGACATCCTTGATTTGCTCATAATACACTCATTCCTTTCACCATTGAGTGT
GCACATATTTCTCTGATTGGAAAGAACTACAGAGGAGGTTTTACNTTTTA
CTTTCCAGTTTGCTATTAAAGAGAGAAAACTAACAGAGNGAAATCAAGCA
ACTCAAAACAACCTTACACACACACACACACACACACACACTCACAAAGATA
TTTTGTTCACCATATGTATTGATGTGCCTGTTTTCAGTCAATCCACAGGAA
GGGCTAAGGAGAGTGACATCTGGGCTACATTAAAAGGACAGTCACATTG
CTCAAAGNACTCAAGTTTAGCCCGAGTACAGTAGCT
Genbank ID: Z24212
Description: H. sapiens (D18S477) DNA segment containing (CA) repeat; clone
Search for GDB entry

27. D18S979:

Database ID: GATA-P28080 (Also known as G08015, CHLC.GATA92C08, CHLC.GATA92C08.P28080)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AGCTTGCAGATAGCCTGCTA
Right = TACGGTAGGTAGGTAGATAGATTCG
Product Length = 155
Review complete sequence:
CTCTACAGTCTCTNACCTTTGGACTCCAGGACTTTCACCAGCACCCTCAA
CATTCCCACTGGGTTCTCAGGACTTTATAGTTGTACTGAGCCATGCCACT
GGATCCTAGGGTCTCCAGCTTGCAGATAGCCTGCTATGGGACTTAATCT
TTGTAATAAGGTGAGTCAATTCTGCCAATAAACCTACTTTCATCTCTATCT
ATCTATCTATCTATCTATCTATCTATCTATATCTATCATCTATCTATCGAAT
CTATCTACCTACCTACCGTATTAGTTCTGTCTCTCTGGAGN

Genbank ID: G08015
Description: human STS CHLC.GATA92C08.P28080 clone GATA92C08.

28. WI-9340:

Database ID: UTR-05134 (Also known as G06102, D18S1034, 9340, X60221)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = TGAGAGAACGAAATCTCTATCGG
Right = AGGCAGCAAGTTTTTATAAAGGC
Product Length = 115
Review complete sequence:

ATGTATCTATCCCAATTGAGTCAGCTAGAAACAGTTGACTGACTAAATGG
AAACTAGTCTATTTGACAAAGTCTTTCTGTGTTGGTGTCTACTGAAGTTAT
AGTTTACCCTTCCTAAAAATGAAAAGTTTGTTTCATATAGTGAGAGAACGA
AATCTCTATCGGCCAGTCAGATGTTTCTCATCCTTCTTGCTCTGCCTTTG
AGTTGTTCCGTGATCATTCTGAATAAGCATTTGCCTTTATAAAAACTTGCT
GCCTGACTAAAGATTAACAGGTTATAGTTTAAATTTGTAATTAATTCTACC
ATCTTGCAATAAAGTGACAATTGAATG

Genbank ID: G06102
Description: WICGR: Random genome wide STSs
Search for GDB entry

29. D18S466:

Database ID: AFM094YE5 (Also known as 094ye5, Z23354, D18S466)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ACACTGTAGCAGAGGCTTGACC
Right = AGGCCAAGTTATGTGCCACC
Product Length = 214
Review complete sequence:
aaatgacactttaaggaggtaacactgtagcagaggcttgaccaccacccagttctcactagcactgagg
atgctctattggttgggttacccacacacgcatagacatgcacacacacagacacacagacacacacac
acacacacacacaccagatatagcattccaaaccatcaatatgctatgcaatactgcattaacaggtcatg
cctgtggtggcacataacttggcctagaaaatactggggacgtctgcattccctttattatcgaattgacttact
tggcttctgagttttcctcagaagtaatacttcaatacctcttccatttctgccttgancattgtttggggtaccaag
tatagct

Genbank ID: Z23354
Description: H. sapiens (D18S466) DNA segment containing (CA) repeat; clone
Search for GDB entry

30. D18S1092:

Database ID: AFMA112WE9 (Also known as D18S1092, w5374, a112we9)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = CTCTCAAAGTAAGAGCGATGTTGTA
Right = CCGAAGTAGAAAATCTTGGCA
Product Length = 163
Review complete sequence:

agctctcaaagtaagagcgatgttgtaactgactgagttgttttgtgaantttgnttttggagtcagtggagcat
gttattagatgtaaatttaaacacacacacacacacacacacacacacacacgagaagtaagtgccaag
attttctacttcggcgcctatatttctatatactgattttctgtatttcccagacttgaatatagattgtctttctgntttat
catagacaatctcataataanttaggcataataaggtaatgaggnttttctgggcttcttttcatcatccctgca
atttgagtctcntttatagntgaantcttctctgtaataacntcttgttttagct

Search for GDB entry

31. D18S61:

Database ID: AFM193YF8 (Also known as 193yf8, Z16834, D18S61)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATTTCTAAGAGGACTCCCAAACT
Right = ATATTTTGAAACTCAGGAGCAT
Product Length = 174
Review complete sequence:
CGTCTTACCAAACCAACATAATATAGCAATGGNAACCAAAAATTTCTAAGA
GGACTCCCAAACTACATTCTTCTNCCTGAATTAAATACAGGCATTCAANA
NAAACANACACACACACACACACACACACACACACACACACACACACGCACA
CCCTTCAAATCNTAGCATAAATTCCNCTTATATAAACATAACCATGCTCCT
GAGTTTCAAAATATTGGGTGGTTCGAAGTTCGAAGCAACAAATTTCCAGT
TAGTGTCTATTANTTGTTGGACAGCT

Genbank ID: Z16834
Description: H. sapiens (D18S61) DNA segment containing (CA) repeat; clone
Search for GDB entry

Markers (STRs) used in refining the candidate region.

Below the markers are shown that were used in family MAD31 to refine the candidate region. Most of these markers are already described above and will therefore only be mentioned to by their name. For the additional markers, the information is given here.

Data was already shown for: D18S68, D18S55, D18S969, D18S1113, D18S483, D18S465, D18S876, D18S477, D18S979, D18S466 and D18S61.

New data:

1. D18S51:

Other names: UT574, (D18S379)
Primer sequences:
UT574a GAGCCATGTTCATGCCACTG
UT574b CAAACCCGACTACCAGCAAC

DNA-sequence:
AATTGAGCNCAGGAGTTTAAGACCAGCCTGGGTAACACAGTGAGACCCC
TGTCTCTACAAAAAAATACAAAAATNAGTTGGGCATGGTGGCACGTGCCT
GTAGTCTCAGCTACTTGCAGGGCTGAGGCAGGAGGAGTTCTTGAGCCCA
GAAGGTTAAGGCTGCAGTGAGCCATGTTCATGCCACTGCACTTCACTCT
GAGTGACAAATTGAGACCTTGTCTCAGAAAGAAAGAAAGAAAGAAAGAAA
GAAAGAAAGAAAGAANGAAAGAAAGAAAGTAAGAAAAAGAGAGGGAAAG
AAAGAGAAANAGNAAANAAATAGTAGCAACTGTTATTGTAAGACATCTCC
ACACACCAGAGAAGTTAATTTTAATTTTAACATGTTAAGAACAGAGAGAAG
CCAACATGTCCACCTTAGGCTGACGGTTTGTTTATTTGTGTTGTTGCTGG
TAGTCGGGTTTGTTATTTTTAAAGTAGCTTATCCAATACTTCATTAACAAT
TTCAGTAAGTTATTTCATCTTTCAACATAAATACGNACAAGGATTTCTTCT
GGTCAAGACCAAACTAATATTAGTCCATAGTAGGAGCTAATACTATCACA
TTTACTAAGTATTCTATTTGCAATTTGACTGTAGCCCATAGCCTTTTGTCG
GCTAAAGTGAGCTTAATGCTGATCGACTCTAGAG

GENBANK ID: L18333

2. D18S346.

Other name: UT575

Primer Pairs:

Primer A: TGGAGGTTGCAATGAGCTG
Primer B: CATGCACACCTAATTGGCG

DNA sequence:
ACGAGGACAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCGTT
TNTACTAAAANTACAAAANTTGGTCGGGAGGCTGGGGCAGGNGACATGC

TTGACCCCAGGAGGTGGAGGTTGCAATGAGCTGAGATTGCACCACTGCA
CTNCAGCNTGG.....AAGAAAGAGAAAGGANAGNNAGGNAGNNANNAAAC
TACATNTGAAGTCAACACTAGTATTGGTGGGAGAGGAATTTTATGCTGCA
TTCCCCNACAACCACTAGATACGCCAATTAGGTGTGCATGGTCCATGCTA
T

GenBank ID: L26588

3. D18S817.

Other name: UT6365

Primer Pairs:

Primer A: GCAAAGCAGAAGTGAGCATG
Primer B: TAGGACTACAGGCGTGTGC

DNA Sequence:

CATATGGGTCCACAAGCAACCTCAGTCCTTGTCTCTTCAGAAGAAAGAAT
TCTACTGAGGGNCATAAGGCAGAAGGAGAGACCTAGGCAAGTTGCAAAG
CAGAAGTGAGCATGTATTAAAAAAGCTTTAGAACAGTAAGGAAAGGAAGAA
AAGAAAAGAAGGAAAGTTCAACTTGGAAGAGGGCCAAGCCGGCAACTTG
GCAGAAGGATTGCTTGAGCCCAGGAGTTAAGACCAGTCTGGGCAATATA
GTGAGACTCCATCTCTGCATACATACATACATACATACATACATACATACA
TACATACATATTGCAGGGTATGATGGCACACGCCTGTAGTCCTAGCTACT
CTGGAGGTTGAGATGGGAGGGTCACTGAGCCTGGGAANTTGAGGCTGC
NNTGAGCCATGATC

GenBank ID: L30552

Characterisation of YACs.

8 YACs were selected covering the candidate region and flanking the gap. These YACs were further characterised by determining the end-sequences by the Inverse-PCR protocol.
Selected YACs: 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752_g_8, 717_d_3, 745_d_2

New STSs based on end-sequences (unless indicated otherwise, the STSs were tested on a monochromosomal mapping pannel for identifying chimaerism of the YAC; if the STS revealed a hit not on chromosome 18q - chimaeric YAC- then it is indicated in the text below):

1. SV32L.

Derived from YAC 745_d_2 left arm end-sequence.

Primer A: GTTATTACAATGTCACCCTCATT
Primer B: ACATCTGTAAGAGCTTCACAAACA

DNA-sequence:

ATTCCTTNGTTATTACAATGTCACCCTCATTTAAAAAGTGGAAAGATAAAG
AGGAAGCAATCTATTTTTTTCCTTTTTTTCTGATAGCACTTGTTTGTGAAG
CTCTTACAGATGTTCTTAAGTAAAATCAACTCCTCCATTTTTTTTGTAGCA
ACTACACATATTTATCAATAATAGTTCACAAATACATTTTCAAATT

Amplified sequence length: 107 basepairs (bp)

This STS has no clear hit on the monochromosomal mapping pannel.

2. SV32R.

Derived from YAC 745_d_2 right arm end-sequence.

Primer A: ACGTTTCTCAATTGTTTAGTC
Primer B: TGTCTTGGCATTATTTTTAC

DNA sequence:

AGACAATGGGAGAAATTGCACTGCCCTGAGTCAGAAATCAGATCTGTTG
CCATACAGCTGCCGTTATGTGATCATTTGCAAGTCAACGTTTCTCAATTG
TTTAGTCATTTGTAAGACAAAAAGACTGGTTGGATTTCAGAGAATTTGGA
ATCCTCCTTCAGGTTTAACAAGCAATAAATGATACTCTTCAGTGTAAAAAT
AATGCCAAGACATNATTTGACTTTAAATTAAATCCAAACAAGATATC

Amplified sequence length: 127 bp

This STS has no clear hit on the monochromosomal mapping pannel.

3. SV11L.

Derived from YAC 766_f_12 left arm end-sequence.

Primer A: CTATGCTCTGATCTTTGTTACTTT
Primer B: ATTAACGGGAAAGAATGGTAT

DNA sequence:

GTCTTTATTTCATATAACTATGCTCTGATCTTTGTTACTTTCTCCTTTTAAC
TCAGTTTAAGCTTTATTCTTATTTTCCAGCTGCTGAAGGTATATAGTTAGG
TTGTTTATTGGATACCATTCTTTCCCGTTAATGTCAGTGGTTACTGCTATC
AATGTAGCAGTTA

Amplified sequence length: 118 bp

This STS has a hit with chromosome 18 and must ba located between CHLC.GATA-p6051 and D18S968.

4. SV11R.

Derived from YAC 766_f_12 right arm end-sequence.

Primer A: AAGGTATATTATTTGTGTCG
Primer B: AAACTTTTCTTAACCTCATA

DNA sequence:

ATAAGGTATATTATTTGTGTCGTGAGTTAAGAAATCATTAATAACTATTTT
CAGAATGACAAATGTCATTATATGTTGTAAAAAAGATAAATACGTGAAATT
ATGAGGTTAAGAAAAGTTTA

Amplified sequence length: 119 bp.

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01.

5. SV34L.

Derived from YAC 717_d_3 left arm end-sequence.

Primer A: TCTACACATATGGGAAAGCAGGAA
Primer B: GCTGGTGGTTTTGGAGGTAGG

ACATAAAATGTCGCTCAAAAACAATTATGTGTGTCTACACATATGGGAAA
GCAGGAAACAAATTTGTTTACAACATACATTACTTTTGTTTTTTAGGCAAG
ATAAAATNTCCTACCTCCAAAACCACCAGCACNGTCCGCAATAACTATAC
ATC

Amplified sequence length: 98 bp

This STS has a hit with chromosome 18.

6. SV34R.

Derived from YAC 717_d_3 right arm end-sequence.

Primer A: ATAAGAGACCAGAATGTGATA
Primer B: TCTTTGGAGGAGGGTAGTC

DNA-sequence:

AATATCATTCTTCACCCACGTTATACATAAGAGACCAGAATGTGATATTGT
CATCTCACATGGAAAAATCTGCTGTGATCAGTTCCTGAAGCTTGCTGTGA
TCCTCCCTTAGGAAAGTAGAAAAATCTTTTTGAAACACTTTATTCTACAAT
CAATGAAAATTAGGTGAAGCTACAGAAGCCAGAAATTACTCTAAGATTAG
ACAATTATTTAAGANGACCAATTGTCTTTGGTCTTCTTCTGAAGGGTCTG
ACTACCCTCCTCCAAAGAATTCACTGGCCGTCGTTTTACAACGTCNTGA

Amplified sequence length: 244 bp

This STS has a hit with chromosome 1, therefore YAC 717_d_3 is chimaeric

7. SV25L.

Derived from YAC 731_c_7 left arm end-sequence.

Primer A: AAATCTCTTAAGCTCATGCTAGTG
Primer B: CCTGCCTACCAGCCTGTC

DNA sequence:

AGTGGAGAGATAGAAAGAGAGGAAGATTTTTTTTTTTTAAATCTCTTAAGCT
CATGCTAGTGTAGGTGCTGGCAGGTCTGAACACTCTGTAGGACAGGCTG
GTAGGCAGGAA

Amplified sequence length: 72 bp

This STS has no clear hits on the monochromosomal mapping pannel.

8. SV25R.

Derived from YAC 731_c_7 right arm end-sequence.

Primer A: TGGGGTGCGCTGTGTTGT
Primer B: GAGATTTCATGCATTCCTGTAAGA

DNA-sequence:

GGAGGGTGTTNTCACANAAGTCTGGGGTGCGCTGTGTTGTTCATTGTAA
AAACCCTTTGGANCATCTGGGAATGTGCTGCCCCACATGTCCAGGTAAC
GTTCTCAGGAAGGGGAGGCTGGAAATCTCTGTGTGTTCTTACAGGAATG
CATGAAATCTCCCANCCCCTCTTGTTGGAAATTTCCCTCACTTT

Amplified sequence length: 136 bp

This STS has a hit with chromosome 7; therefore YAC 731_c_7 is chimaeric

9. SV31L.

Derived from YAC 752_g_8 left arm end-sequence.

Primer A: GAGGCACAGCTTACCAGTTCA
Primer B: ATTCATTTTCTCATTTTATCC

DNA-sequence:

CTTCTCNATGANTGGACAAATGTCATTGGGTCAGCATGAGGCACAGCTT
ACCAGTTCAGATTCCAGTAGCTGAGGAACAAATCTTAACTCCAAAAATAA
GTAATTGCGTCACTTTGGAGGAATTATTTGACCTTTTCATAACTTTGACAT
CACAACAATGAGGGTGAAGTTAGTAAAATAAATGATTATTATGAGGATAA
AATGAGAAAATGAATTNAGTGCTTAAGACAATGCTTGGTAACTAGTTAAN
CCG

Amplified sequence length: 178 bp

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01.

10. SV31R.

Derived from YAC 752_g_8 right arm end-sequence.

Primer A: CAAGATTATGCCTCAACT
Primer B: TAAGCTCATAATCTCTGGA

DNA sequence:

AAACTTTAACCAATTTAAACTCCCTAACAGTTCTATAAAATAAGCAAGATT
ATGCCTCAACTTTATGGATAAAGAAATGGAGGCATTAAGAGATAACTAAC
TTGCCCAAGGCCACACAAGTGACTGAGTAAGAATTGCAAAGCCAATGAG
TCTGGCTCCAGAGATTATGAGCTTAATCACCACACTGTGCCACCTCCTGT
GTTTCCTGG

Amplified sequence length: 131 bp

This STS has no clear hits on the monochromosomal mapping pannel and gives no information concerning the chimaerity of the YAC.

11. SV10L.

Derived from YAC 942_c_3 left arm end-sequence.

Primer A: TCACTTGGTTGGTTAACATTACT
Primer B: TAGAAAAACAGTTGCATTTGATAT

DNA-sequence:

GGTNTTTCACTTGGTTGGTTAACATTACTTCTAAGTTTTTTATTGTTTTTTA
TGCTATTGCTAATGGGATTGCTTTCTTAATTTATTTTTTCCAATAGCTTGT
TGTTAGTTTATATCAAATGCAACTGTTTTTCTATGCAAATTATGTTTCCT

Amplified sequence length: 130 bp

This STS has a hit with chromosome 18 and must be located between CHLC.GATA-p6051 and D18S968

12. SV10R.

Derived from YAC 942_c_3 right arm end-sequence.

Primer A: AACCCAAGGGAGCACAACTG
Primer B: GGCAATAGGCTTTCCAACAT

DNA sequence:
TTGGTGGTGCCCTAGGTTTGGCAATTATAAATAAAGCTGCTACAAACATT
CATGTGCAGGTCTCCGTGTGGACATAATTTTCCAGTTCATTTGGGTAAAA
CCCAAGGGAGCACAACTGTTGGATCCTATNATAAAAATATNTCTCGTTTC
ATTTAAAAAACCTGGGAAACTATCTNCCCACAGTGGCTGTCCCTTTTTGT
ATCCCCACCAACAATGTTGGAAAGCCTATTGCCANCAT

Amplified sequence length: 135 bp

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01

13. SV6L.

Derived from YAC 961_h_9 left arm end-sequence.

No primer was made, because this sequence is identical to a known STR marker D18S42, which is indeed mapped to this region.

Primer A:
Primer B:

DNA sequence:

CATGNCTCACAGTGTTCTGAGGCTGCTCTGGACATGCAATCTTGCATGC
TTTTGTCATGACAGGTCTTAAANAGTTTATCAGCTTNCTCAAATAGCTGAA
TGACANAACACTGGATTTTTGTTCAAATANCCTATCAACTTGGCNTCTGT
GTTGCGGTTGTCACTTGGTAACAAAATAAGTC

Amplified sequence length:

SV6L recognises D18S42 which must be therefore located between WI-7336 and WI-8145

14. SV6R.

Derived from YAC 961_h_9 right arm end-sequence.

Primer A: TTGTGGAATGGCTAAGT
Primer B: GAAAGTATCAAGGCAGTG

DNA sequence:

TAATTGACAAATAAAAATTGTATATTTTNCATATTTAACATGTTATGCTAAC
ATATATATGGATTGTGGAATGGCTAAGTCAGAAATTCTTTTACATTCATAT
TTCCATATTATTTACTTTNNGCTTTAAAAAATATGTAAATGANAATACTTAT
TTTTTTCAGTGTCACTGCCTTGATACTTTCACATTTNNGTTACATATTATTT
CCCTTNCATCTAACAAATATATATTGAGTTTCTATAATGTGTCTGACACTG
A

Amplified sequence length: 122 bp

SV6R amplifies a segment on chromosome 18. This segment must be located between WI-2620 and WI-4211

15. SV26L.

Derived from YAC 907_e_1 left arm end-sequence.

Primer A: TATTTGGTTTGTTTGCTGAGGT
Primer B: CAAGAAGGATGGATACAAACAAG

DNA sequence:

TGGTCACTGGTGCCTTATTTGGTTTGTTTGCTGAGGTCATATTTCCTGTG
GCCTTCATGCTTGATTTGTTGGAGTCTAGCCATGTAAAANTCTGTTGGAG
TCTAGGCATTTAAAAAATAGGTATTTATTGTAATCTTTGCCATTTGCTTGT
TTGTATCCATCCTTCTTGGGAAGGCTTTACAGGCATTCAAAAGG

Amplified sequence length: 154 bp

This STS has a hit with chromosome 13; therefore YAC 907_e_1 is chimaeric.

16. SV26R.

Derived from YAC 907_e_1 right arm end-sequence.

Primer A: CGCTATGCATGGATTTA
Primer B: GCTGAATTTAGGATGTAA

DNA sequence:

CGCTATGCATGGATTTAAACTGAGTGTAGTGCACTCACTATGTTGCAGTC
TCTTATTCTAGGTTCCTAATATTTACATCCTAAATTCAGCT

Amplified sequence length: 90 bp

no clear hits on monochromosomal mapping pannel: no information concerning chaemerity at this side of the YAC

Testing of 3 end-sequences flanking the gap in additional YACs: STS-markers WI-4211, D18S876 and GCT3G01 are also shown in order to identify YACs on opposite sides of the gap more clearly in table 3 below.

	STSs					
YACs	WI-4211	D18S876	SV31L	SV11R	SV10R	GCT3G01
940_b_1	+	+	+	-	-	-
766_f_12	+	+	+	+	-	-
846_a_5	+	-?	+	+	-	-
752_g_8	+	+	+	+	-	-
745_d_2	+	+	+	+	-	-
961_c_1	+	+	-	-	-	-
942_c_3	+	+	+	+	+	-
717_d_3	-	-	+	+	-?	+
972_e_11	-	-	-	-	-	+
940_h_10	-	-	-	-	+	+
821_e_7	-	-	-	-	+	+
731_c_7	-	-	-	-	-	+
889_c_4	-	-	-	-	+	+
907_e_1	-	-	-	+	+	+

- +: positive hit / -: no hit / ?: 2 instances were observed in which a positive hit was expected (on the assumed order of the markers) but not observed. The reasons for this are not clear.

YAC 745_d_2 was excluded from further analysis since there was no clear hit with chromosome 18. Of the remaining 7 from a monochromosomal mapping panel it was determined that 3 were chimeric and 4 non-chimeric as shown in Table 4 below.

TABLE 4

YAC	chimaeric	chromosome
961_h_9 (6)	no	
942_c_3 (10)	no	
766_f_12 (11)	no	
731_c_7 (25)	yes	chromosome 7
907_e_1 (26)	yes	chromosome 13
752_g_8 (31)	no	
717_d_3 (34)	yes	chromosome 1

For the non-chimeric YACs the STS based on the end-sedquence flanking the gap (10R, 11R, 31L) was tested on 14 YACs flanking the gap. Overlaps between YACs on opposite sides of the gap were demonstrated: e.g. the "11R" end-sequence (766.f.12) detects YAC 766.f.12 and YAC 907.e.1.
YACs were then selected comprising the minimum tiling path:

TABLE 5

YAC	size	chimaerity
961_h_9	1180 kb	not chimaeric
766_f_12	1620 kb	not chimaeric
907_e_1	1690 kb	chimaeric (chr. 13)

These three YACs are stable as determined by PFGE and their sizes roughly correspond to the published sizes. These YACs were transferred to other host-yeast strains for restriction mapping.

Experimental 2

Construction of fragmentation vector:

A 4.5kb ECORI/SalI fragment of pBLC8.1 (Lewis *et al*, 1992) carrying a lysine-2 and a telomere sequence was directionally cloned into GEM3zf(-) digested with ECORI/SalI. Subsequently, an End Rescue Site was ligated into the EcoRI site. Hereto, two oligonucleotides (strand 1: 5'-TTCGGATCCGGTACCATCGAT-3' AND STRAND 2: 3'-GCCTAGGCCATGGTAGCTATT-5') were ligated into a partial (dATP) filled ECORI site, generating the vector pDF1. Triplet repeat containing fragmentation vectors were constructed by cloning of a 21bp and a 30bp CAG/CTG adapter into the Klenow-filled PstI site of pDF1. Trasformation and selection resulted in a $(CAG)_7$ and a $(CTG)_{10}$ fragmentation vector with the orientation of the repeat sequence 5' to 3' relative to the telomere.

Yeast transformation:

Linearised (digested with SalI) vector was used to transform YAC clones 961.h.9, 766.f.12 or 907.e.1 using the LiAc method. After transformation the YAC clones were plated onto SDLys$^-$ plates to select for the presence of the fragmentatio vector. After 2-3 days colonies were replica plated onto SDLys$^-$-Trp$^-$-Ura$^-$ and SDLys$^-$-Trp$^-$-Ura$^+$ plates. Colonies growing on the SDLys$^-$-Trp$^-$-Ura$^+$ plates but not on the SDLys$^-$-Trp$^-$-Ura$^-$ plates contained the fragmented YACs.

Analysis of fragmented YACs:

Yeast DNA isolated from clones with the correct

phenotype was analysed by Pulsed Field Electrophoresis (PFGE), followed by blotting and hybridisation with the Lys-2 gene and the sizes of the fragmented YACs were estimated by comparison with DNA standards of known length.

End Rescue:

Fragmented YACs characterised by a size common to other fragmented YACs, indicative of the presence of a major CAG or CTG triplet repeat, were digested with one of the enzymes from the End Rescue site, ligated and used to transform E. Coli. After growth of the transformed bacteria the plasmid DNA was isolated and the ends of the fragmented YACs, corresponding to one of the sequences flanking the isolated trinucleotide repeats, were sequenced.

Sequencing revealed that fragmented YACs of an equal length were all fragmented at the same site. A BLAST Search of the GenBank database was performed with the identified sequences to identify homology with known sequences. The complete sequence spanning the CAG or CTG repeats of the fragmented YACs was obtained by Cosmid Sequencing, employing sequence specific primers and splice primers, as previously described (Fuentes *et al.* 1992 Hum.Genet. **101**: 346-350) or by using the "genome walker" kit (Clontech Laboratories, Palo Alto, USA) and described in Siebert *et al.* Nucleic Acid Res (1995) **23**(6): 1087-1088 and Siebert *et al.* (1995) CLONTECHniques **X**(II): 1-3.

Results:

A YAC 961_h_9 clone was transformed with the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. The CTG vector

did not reveal the presence of any CTG repeat. Analysis of twelve $(CAG)_7$ fragmented YACs showed that five of these had the same size of approximately 100kb. End Rescue was performed with ECORI and sequencing of three of these fragments revealed that they all shared the terminal sequence shown in italics in Figure 15a. A BLAST search of the Genbank database with this sequence indicated the presence of a sequence homology with the CAP2 gene (GenbBank accession number: L40377). The sequence spanning the CAG repeat shown in Figure 15a was obtained by both cosmid sequencing and genome walker sequencing. The sequence was mapped between markers D18S68 and WI-3170 by STS content mapping.

A YAC 766-f-12 was fragmented using the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. Again the $(CTG)_{10}$ vector did not reveal the presence of any CTG repeat. Analysis of twenty $(CAG)_7$ fragmented YACs showed the presence of two groups of fragments with the same size: five of approximatively 650kb and two of approximatively 50kb.

End Rescue was performed using ECORI on four of the fragmented YACs of 650kb. Sequencing confirmed that they all shared identical 3' terminals, characterised by the sequence shown in italics in Figure 16a. A Blast Search showed homology of this sequence with the Alu repeat sequence family. The sequence spanning the CAG repeat shown in Figure 16a was obtained by cosmid sequencing. The sequence was mapped between markers WI-2620 and WI-4211 by STS content mapping on the YAC contig map.
End Rescue was also performed on the two fragments of 50kb. Sequencing revealed the sequence shown in italics in figure 17a. A Blast Search revealed no

sequence homology with any known sequence. Cosmid sequencing allowed to identify the complete sequence spanning the CAG repeats, shown in figure 17a. The sequence was mapped between markers D18S968 and D18S875 by STS content mapping on the YAC contig map.

A YAC 907-e-1 clone was transformed with the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. The $(CAG)_7$ vector did not reveal the presence of any CAG repeat. Analysis of twenty-six $(CTG)_{10}$ fragmented YACs revealed that twenty-one of them had the same size of approximatively 900kb. End Rescue was performed with KpnI on three fragmented YACS of this size. Sequencing revealed the nucleotide sequence shown in italics in Figure 18a. A Blast Search indicated the presence of an homology of this sequence with the GCT3G0I marker (GenBank accession number: G09484). The sequence spanning the CTG repeat was obtained from the GenBank Database. The sequence was mapped between markers 10R and WI-528.

CLAIMS:

1. Use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder.

2. Use of a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61 for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder.

3. The use as claimed in claim 2 wherein said portion comprises the region of chromosome 18q between polymorphic markers D18S68 and D18S979 or a fragment of said region.

4. The use as claimed in claim 2 or 3 wherein said YAC clone is 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752-g-8 or 717_d_3.

5. The use as claimed in claim 4 wherein said YAC clone is 961_h_9, 766_f_12 or 907_e_1.

6. The use as claimed in any preceding claim wherein said mood disorder or related disorder is selected from the Diagnostic and Statistical Manual of Mental Disorders, version 4 (DSM-IV) taxonomy and includes mood disorders (296.XX, 300.4, 311, 301, 13, 295.70), schizophrenia and related disorders (295, 297.1, 298.9, 297.3, 298.9), anxiety disorders

(300.XX, 309.81, 308.3), adjustment disorders (309, XX) and personality disorders (codes 301. XX).

7. A method of identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which comprises detecting nucleotide triplet repeats in a region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

8. A method of identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which comprises fragmentation of a YAC clone as defined in any one of claims 2 to 4 and detection of nucleotide triplet repeats.

9. A method as claimed in claim 7 or 8 wherein said repeated triplet is CAG or CTG.

10. A method as claimed in claim 9 wherein said repeated triplet is detected by means of a probe comprising at least 5 CTG and/or CAG repeats.

11. A method of identifying at least one human gene including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder wherein said gene is present in the DNA comprised in the YAC clones as defined in any one of claims 2 to 5, which method comprises the step of detecting an expression product of said gene with an antibody capable of recognising a protein with an amino acid sequence comprising a string of at least 8 continuous glutamine residues.

12. A method as claimed in claim 11 wherein said DNA forms part of a human cDNA expression library.

13. A method as claimed in claim 11 or claim 12 wherein said antibody is mAB 1C2.

14. A method of preparing a contig map of YAC clones of the region of human chromosome 18q between polymorphic markers D18S60 and D18S61 which comprises the steps of:

(a) subcloning the YAC clones according to any one of claims 2 to 5 into exon trap vectors;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other known sequence tagged sequence from the YAC contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to detect overlaps among the cosmid vectors, and

(c) constructing a cosmid contig map of a YAC clone of said region.

15. A method of identifying at least one human gene or any mutated or polymorphic variant thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) transfecting mammalian cells with DNA sequences cloned into an exon trap vector as prepared in claim 14;

(b) culturing said mammalian cells in an appropriate medium;

(c) isolating RNA transcripts expressed from an SV40 promoter;

(d) preparing cDNA from said RNA transcripts;

(e) identifying splicing events involving exons of the DNA subcloned into said exon trap vector in accordance with claim 14 to elucidate positions of coding regions in said subcloned DNA;

(f) detecting differences between said coding regions and equivalent regions in the DNA of an individual afflicted with said mood disorder or related disorder; and

(g) identifying said gene or mutated or polymorphic variants thereof which is associated with said mood disorder or related disorder.

16. A method of identifying at least one human gene or mutated or polymorphic variants thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) subcloning the YAC clones according to any one of claims 2 to 5 into a cosmid, BAC, PAC or other vector;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other known sequence tagged sequence from the YAC contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to defect overlaps amongst the subclones and construct a map

thereof;

(c) identifying the position of genes within the subcloned DNA by one or more of CpG island identification, zoo-blotting, hybridization of said subcloned DNA to a cDNA library or a Northern blot of mRNA from a panel of culture cell lines;

(d) detecting differences between said genes and equivalent regions of the DNA of an individual afflicted with a mood disorder or related disorder; and

(e) identifying said gene which, if defective, is associated with said mood disorder or related disorder.

17. An isolated human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which is obtainable by the method according to any of claims 7 to 13, 15 or 16.

18. A human protein which, if defective, is associated with a mood disorder or related disorder which is the expression product of the gene according to claim 17.

19. A cDNA encoding the protein of claim 18 which is obtainable by the method of any one of claims 7 to 13 , 15 or 16.

20. Use of a probe of at least 14 contiguous nucleotides of the cDNA of claim 19 or the complement thereof in a method for detection in a patient of a

pathological mutation or genetic variation associated with a mood disorder or related disorder which method comprises hybridizing said probe with a sample from said patient and from a control individual.

21. A nucleic acid molecule which comprises a sequence of nucleotides as shown in any one of Figures 15a, 16a, 17a or 18a.

22. A nucleic acid molecule which comprises a sequence of nucleotides which differ from a sequence of nucleotides as shown in any one of Figures 15a, 16a, 17a or 18a only in the extent of trinucleotide repeats.

23. A protein encoded by a nucleic acid molecule as claimed in claim 21.

24. A protein encoded by a nucleic acid molecule as claimed in claim 22.

25. A method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises analysing a sample of DNA from that individual for the presence of a DNA polymorphism associated with a mood disorder or related disorder in a region of chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

26. A method as in claims 25 wherein said DNA polymorphism is a trinucleotide repeat expansion.

27. A method as in claim 26 wherein said trinucleotide repeat expansion is comprised in a

sequence of nucleotides that differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a or 18a only in said trinucleotide repeat expansion.

28. A method as in claim 26 or 27 which comprises the steps of:

a) obtaining a DNA sample from said individual;

b) providing primers suitable for the amplification of a nucleotide sequence comprised in the sequence shown in any one of Figures 15a, 16a, 17a or 18a said primers flanking the trinucleotide repeats comprised in said sequence;

c) applying said primers to the said DNA sample and carrying out an amplification reaction;

d) carrying out the same amplification reaction on a DNA sample from a control individual; and

e) comparing the results of the amplification reaction for the said individual and for the said control individual;

wherein the presence of an amplified fragment from said individual which is bigger in size from that of said control individual is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

29. A method as in claim 28 wherein said

nucleotide sequence to be amplified is comprised in the sequence shown in Figure 15a and said primers have the sequences shown in Figure 15b.

30. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 16a and said primers have the sequences shown in Figure 16b.

31. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 17a and said primers have the sequences shown in Figure 17b.

32. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 18a and said primers have the sequences shown in Figure 18b.

33. A method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises the steps of :

a) obtaining a protein sample from said individual; and

b) detecting the presence of the protein of claim 24;

wherein the presence of said protein is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

34. A method as in claim 33 wherein said protein is detected with an antibody which is capable of

recognising a string of at least 8 continuous glutamines.

35. A method as in claim 34 wherein said antibody is mAB 1C2.

36. A nucleic acid as claimed in claim 21 for use as a medicament in the treatment of a mood disorder or related disorder.

37. A protein as claimed in claim 23 for use as a medicament in the treatment of a mood disorder or related disorder.

38. A pharmaceutical composition which comprises a nucleic acid as claimed in claim 21 and a pharmaceutically acceptable carrier.

39. A pharmaceutical composition which comprises a protein as claimed in claim 23 and a pharmaceutically acceptable carrier.

40. An expression vector which comprises a sequence of nucleotides as claimed in claims 21 or 22.

41. A reporter plasmid which comprises the promoter region of a nucleic acid molecule as claimed in claim 21 or 22 positioned upstream of a reporter gene which encodes a reporter molecule so that expression of said reporter gene is controlled by said promoter region.

42. A cell line transfected with the expression vector of claim 40.

43. An eukaryotic cell or multicellular tissue or organism comprising a transgene encoding a protein as claimed in claims 23 or 24.

44. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell as claimed in claim 42 with said compound;

b) detecting and/or quantitatively evaluating the presence of any mRNA transcript corresponding to a nucleic acid as claimed in claim 21 or 22; and

c) comparing the level of transcription of said nucleic acid with the level of transcription of the same nucleic acid in a cell as claimed in claim 42 not exposed to said compound;

45. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell as claimed in claim 42 with said compound;

b) detecting and/or quantitatively evaluating the expression of a protein as claimed in claims 23 or 24 and

c) comparing the level of expression of said

protein with that of the same protein in a cell not exposed to said compound.

46. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell transfected with a reporter plasmid as claimed in claim 41 with said compound;

b) detecting or quantitatively evaluating the amount of reporter molecule expressed; and

c) comparing said amount with the amount of expression of said reporter molecule in a cell comprising said reporter plasmid and not exposed to said compound.

47. A compound identified as an enhancer or an inhibitor of the expression of a gene associated with a mood disorder or related disorder by a method as claimed in claims 44 to 46.

FIG. 1.

GTCTTTATTTCATATAACTATGCTCTGATCTTTGTTACTTTCTCCTTTTAAC
TCAGTTTAAGCTTTATTCTTATTTTCCAGCTGCTGAAGGTATATAGTTAGG
TTGTTTATTGGATACCATTCTTTCCCGTTAATGTCAGTGGTTACTGCTATC
AATGTAGCAGTTA

FIG. 2.

ATAAGGTATATTATTTGTGTCGTGAGTTAAGAAATCATTAATAACTATTTT
CAGAATGACAAATGTCATTATATGTTGTAAAAAAGATAAATACGTGAAATT
ATGAGGTTAAGAAAAGTTTA

FIG. 3.

ACATAAAATGTCGCTCAAAAACAATTATGTGTGTCTACACATATGGGAAA
GCAGGAAACAAATTTGTTTACAACATACATTACTTTTGTTTTTTAGGCAAG
ATAAAATNTCCTACCTCCAAAACCACCAGCACNGTCCGCAATAACTATAC
ATC

FIG. 4.

AATATCATTCTTCACCCACGTTATACATAAGAGACCAGAATGTGATATTGT
CATCTCACATGGAAAAATCTGCTGTGATCAGTTCCTGAAGCTTGCTGTGA
TCCTCCCTTAGGAAAGTAGAAAAATCTTTTTGAAACACTTTATTCTACAAT
CAATGAAAATTAGGTGAAGCTACAGAAGCCAGAAATTACTCTAAGATTAG
ACAATTATTTAAGANGACCAATTGTCTTTGGTCTTCTTCTGAAGGGTCTG
ACTACCCTCCTCCAAAGAATTCACTGGCCGTCGTTTTACAACGTCNTGA

FIG. 5.

GGAGGGTGTTNTCACANAAGTCTGGGGTGCGCTGTGTTGTTCATTGTAA
AAACCCTTTGGANCATCTGGGAATGTGCTGCCCCACATGTCCAGGTAAC
GTTCTCAGGAAGGGGAGGCTGGAAATCTCTGTGTGTTCTTACAGGAATG
CATGAAATCTCCCANCCCCTCTTGTTGGAAATTTCCCTCACTTT

FIG. 6.

CTTCTCNATGANTGGACAAATGTCATTGGGTCAGCATGAGGCACAGCTT
ACCAGTTCAGATTCCAGTAGCTGAGGAACAAATCTTAACTCCAAAAATAA
GTAATTGCGTCACTTTGGAGGAATTATTTGACCTTTTCATAACTTTGACAT
CACAACAATGAGGGTGAAGTTAGTAAAATAAATGATTATTATGAGGATAA
AATGAGAAAATGAATTNAGTGCTTAAGACAATGCTTGGTAACTAGTTAAN
CCG

FIG. 7.

GGTNTTTCACTTGGTTGGTTAACATTACTTCTAAGTTTTTTATTGTTTTTTA
TGCTATTGCTAATGGGATTGCTTTCTTAATTTATTTTTTCCAATAGCTTGT
TGTTAGTTTATATCAAATGCAACTGTTTTTCTATGCAAATTATGTTTCCT

FIG. 8.

TTGGTGGTGCCCTAGGTTTGGCAATTATAAATAAAGCTGCTACAAACATT
CATGTGCAGGTCTCCGTGTGGACATAATTTTCCAGTTCATTTGGGTAAAA
CCCAAGGGAGCACAACTGTTGGATCCTATNATAAAAATATNTCTCGTTTC
ATTTAAAAAAACCTGGGAAACTATCTNCCCACAGTGGCTGTCCCTTTTTGT
ATCCCCACCAACAATGTTGGAAAGCCTATTGCCANCAT

FIG. 9.

CATGNCTCACAGTGTTCTGAGGCTGCTCTGGACATGCAATCTTGCATGC
TTTTGTCATGACAGGTCTTAAANAGTTTATCAGCTTNCTCAAATAGCTGAA
TGACANAACACTGGATTTTTGTTCAAATANCCTATCAACTTGGCNTCTGT
GTTGCGGTTGTCACTTGGTAACAAAATAAGTC

FIG. 10.

TAATTGACAAATAAAAATTGTATATTTTNCATATTTAACATGTTATGCTAAC
ATATATATGGATTGTGGAATGGCTAAGTCAGAAATTCTTTTACATTCATAT
TTCCATATTATTTACTTTNNGCTTTAAAAAATATGTAAATGANAATACTTAT
TTTTTTCAGTGTCACTGCCTTGATACTTTCACATTTNNGTTACATATTATTT
CCCTTNCATCTAACAAATATATATTGAGTTTCTATAATGTGTCTGACACTG
A

FIG. 11.

TGGTCACTGGTGCCTTATTTGGTTTGTTTGCTGAGGTCATATTTCCTGTG
GCCTTCATGCTTGATTTGTTGGAGTCTAGCCATGTAAAANTCTGTTGGAG
TCTAGGCATTTAAAAAATAGGTATTTATTGTAATCTTTGCCATTTGCTTGT
TTGTATCCATCCTTCTTGGGAAGGCTTTACAGGCATTCAAAAGG

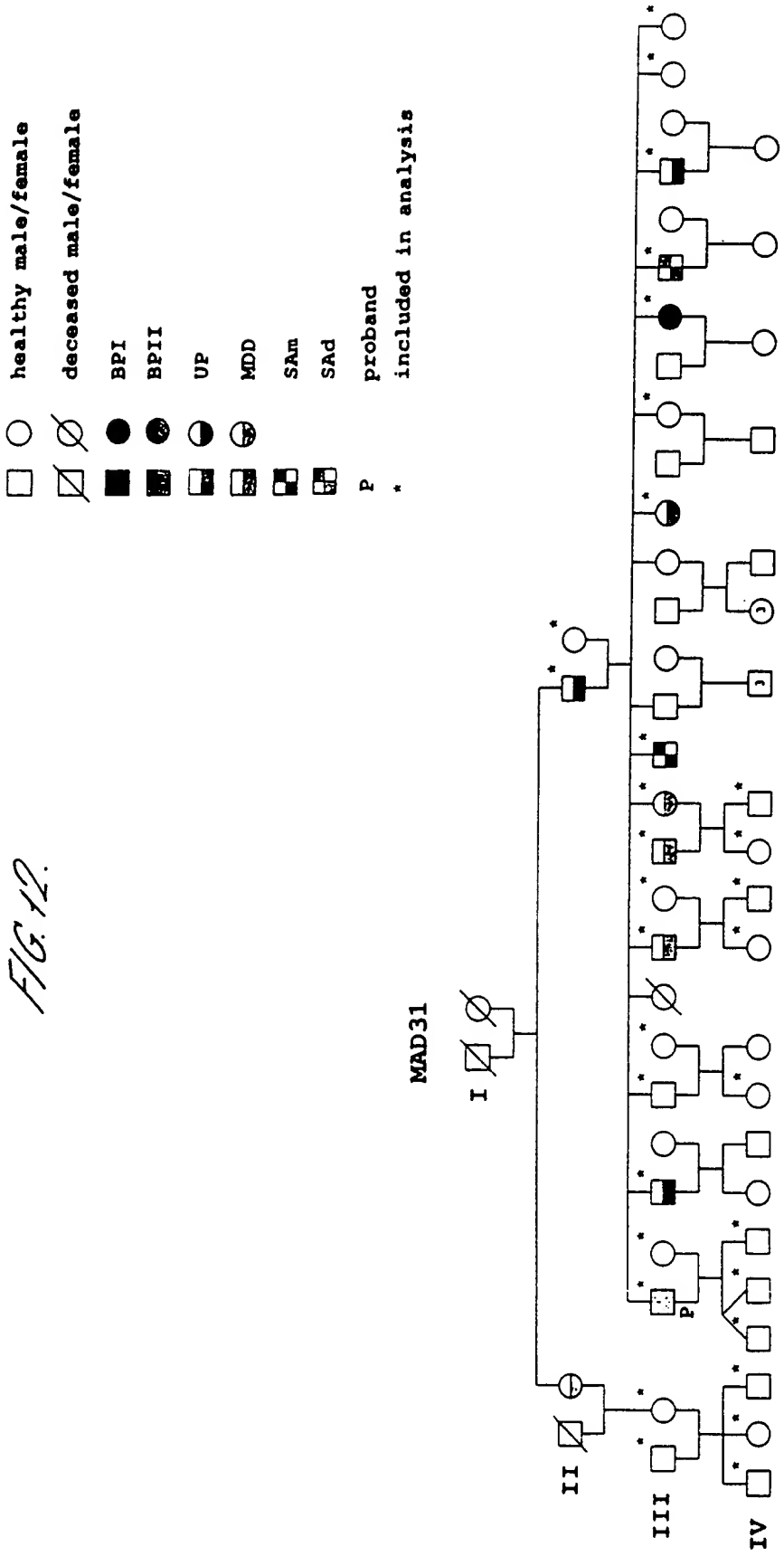
healthy male/female
deceased male/female
BPI
BPII
UP
MDD
SAm
SAd
P
proband
*
included in analysis
MAD31
I
II
III
IV
P

FIG 12.

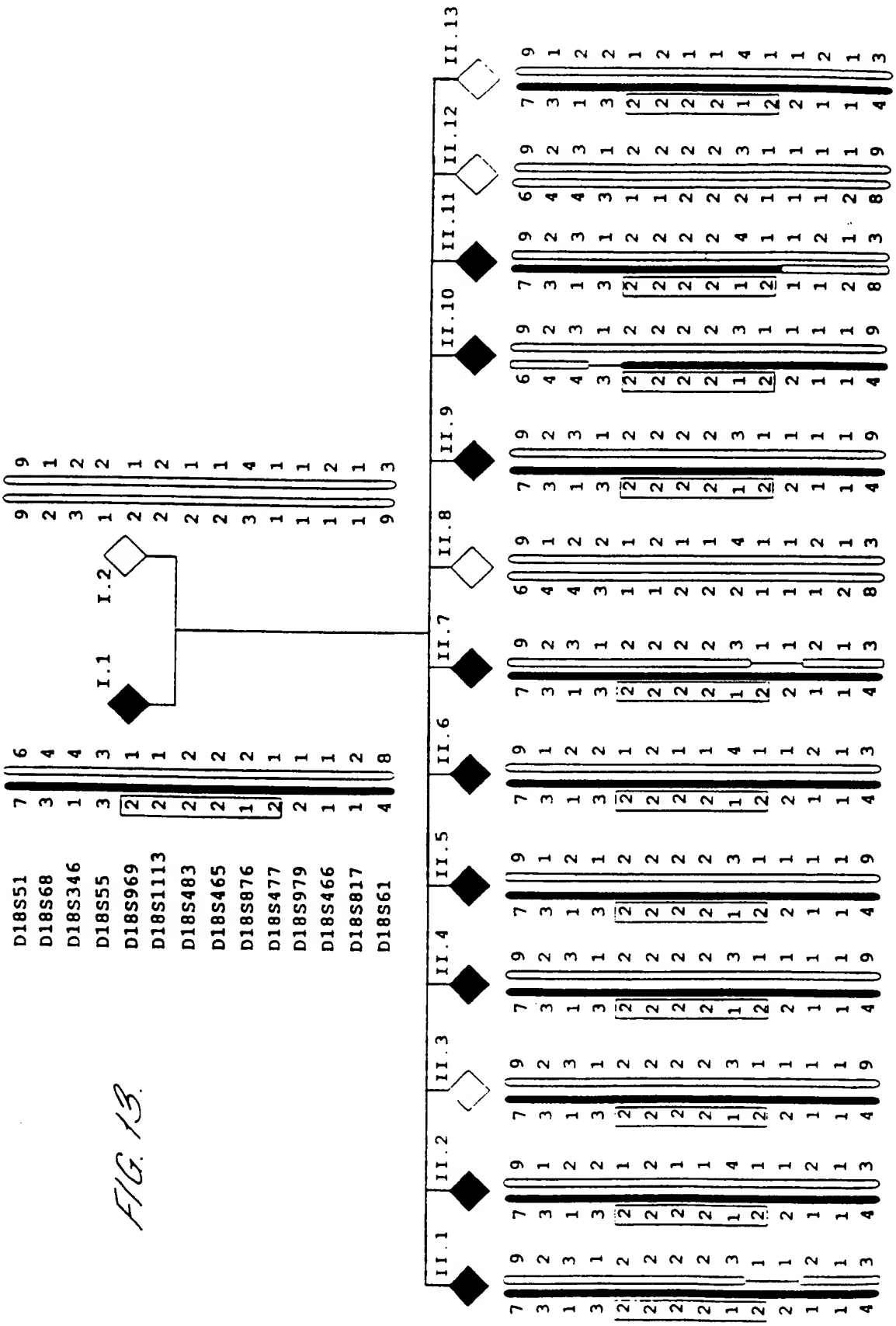
D18S51
D18S68
D18S346
D18S55
D18S969
D18S1113
D18S483
D18S465
D18S876
D18S477
D18S979
D18S466
D18S817
D18S61
I.1
I.2
II.1
II.2
II.3
II.4
II.5
II.6
II.7
II.8
II.9
II.10
II.11
II.12
II.13

FIG 13.

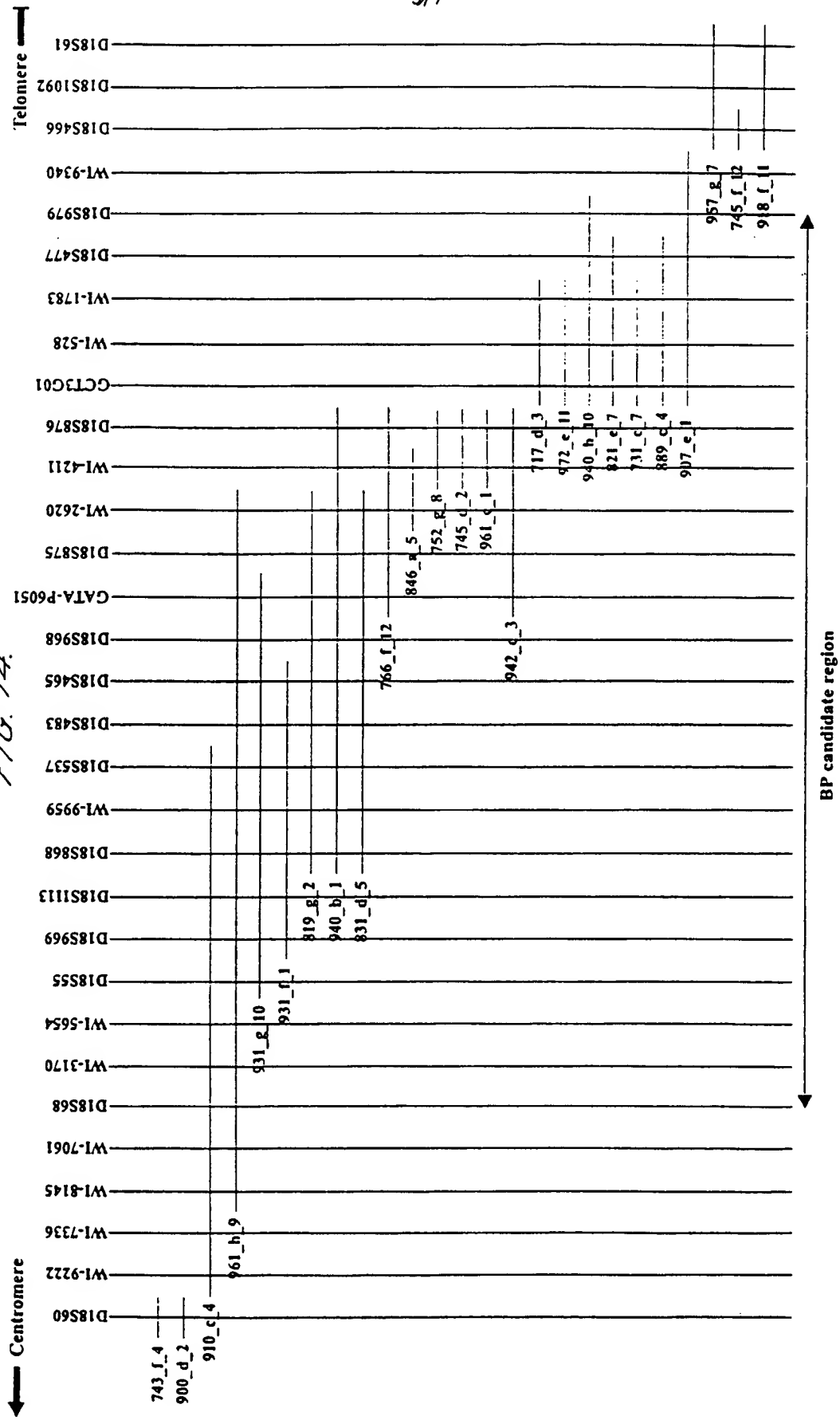
Telomere
D18S61
D18S1092
D18S466
WI-9340
D18S979
D18S477
WI-1783
WI-528
GCT3C01
D18S876
WI-4211
WI-2620
D18S875
GATA-P6051
D18S968
D18S465
D18S483
D18S537
WI-9959
D18S868
D18S1113
D18S969
D18S55
WI-5654
WI-3170
D18S68
WI-7061
WI-8145
WI-7336
WI-9222
D18S60
Centromere
743_f_4
900_d_2
910_c_4
961_h_9
931_g_10
931_f_1
819_g_2
940_b_1
831_d_5
766_f_12
846_g_5
752_g_8
745_d_2
961_c_1
942_c_3
717_d_3
972_e_11
940_h_10
821_e_7
731_c_7
889_c_4
907_e_1
957_g_7
745_f_12
918_f_11
BP candidate region

FIG. 14.

FIG. 15a.

GCCAACAAACAAAATGAAATAAGACCTGGGATGTATTTTTTGGCCAAGGCAATTAGAAA
ATGATTAGTATCCTTATCAGGAGCAATTTCAGAGAATGTTTGGGTGGACGTCTAACTACA
GTGGAGTCAAACGTGAATCAACGGTGAAAAAAGGACAATAGCCAATGTGTACACTTTTT
ATAAAAACCACCCTCCAAGGACCAGGCACTGGCCCTCTCTCCGGTGCCCACAGACATC
CACACAGGCCCAAAGAATCAGGGATTGCACAAGCCAGAGCAATCGAACGGTTCTGAGT
CATCTGCCGGAAGCCTTGCCCTCAATCAAGGCGGACGTGAAGCATCTACAAAGGAGGA
*ATAGTCAAAG***CAGCAG**CGGCGGCGGCGGCGGCGGC**AGCAGCAGCAGCAGCAGG**
AGGTGGGGGCCTCTGCCAGGTACCGGGCGGGGCAGGCACGGAGGTGCCCAGGTT
CCCGCGGAGGCCACCTCTTCCCTGGAGTGCGTGAGAGAGGGGAAGGGAGGAAGG
CCAGAGCAGGAATCAGAGCGAGGCAAAGGCGGGCAGGAACTAXGAGAATGACS
GCGGGAGGCGGCCGGGAAAGAAAXTCTCGGGGCTGTGGGGGTCXCCCTGGCACC
AGCCGGGGTCCCAAGCCCCACCGCGAGACCCCGCGA

FIG. 15b.

5'-ATCGAACGGTTCTGAGTCATCT
5'-CGCTCTGATTCCTGCTCTG

FIG. 16a.

TTCAGTAGAAGGAAGCACAGCAAATTTGCCTTTATAGAGATTCAATTCTTGGTGCTTGG
GCCAAAGAATAAGAATTACATTAAGCAGGCCGGGCACGGTGGCTCACACCTGTAAAAC
CAGAACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAGATCGAGACCATCC
TGGACAACATAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTG
GTGCATGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGCAGGAGAATCCCTTGAACCA
GGGAGTTGGAGGTTGCAGTGAGCCGAGATCACGCCACAGCACTCTAGCCTGGCGACA
*GAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAATTACATTAAG***CAGCAGC**
AGCAGCAGTGASAGAGGGAAKAATGAAAGAAGAAATTTCTAGAATAAGATTGA
TCTCCAGCACCATGCCAATCATGGACTGGATACAATTCATGCATATCTTTTGTGA
GAGAGGTGAGAGATGTGAATCCTTTCTCATT

FIG. 16b.

5'-AGAAGGAAGCACAGCAAATTTG
5'-GCATGGTGCTGGAGATCAAT

FIG. 17a.

TGGGAGTTAAAGCAGACATTCGGCTTTNGTGTTGCCAGAGTTCTAACATAAGTTCTTTTT
CATCTGGGCAGGCNGATGTTCCTTCCATCTTNGAAGNACNGTCCTTTTCATTTTTTTTAT
TTNGCTTTTGGSKTTTATCTTCTTAGACGTCTTCAGGAGTTKGATTGTAGKGTAAGGCAG
ATTTAGTTGACTGGGCTTTGTTTCTGGAAAATTTTAAAGGGGCAAGTCCTGGGCTGCAT
ATTCTTACTCTGGGGGCTTAGTACTGGCCCCTAAATTTGTTCTCTGGCTCCTCAAGGTT
AGAAATCTGCTGGCTGGAGGGGCTGAGATGTTCCTTGACTGCTGGCCAGAACATTCCG
CCGGGGGGTGGCAACCGAAGTGTTTCTTTGGGCAATGGCAGCAGAATTCATGATTGTT
*TTCATGTRC***CAGCAGCAG**TGGCAGCGCAKTGAGTTGCATGATTGTTGGCTGGGGC
TGAGTGCTGGCASGCACTGGAGTGTTTGGCTTCCAGTAGAAATTCACAGCAGTAG
TAGTGGTGGCATGGGAAGGAGGGCAGYGGTGGCATGGGGAGGACCCCCC

FIG. 17b.

5'-GGCTGAGATGTTCCTTGACTGC
5'- CCTTCCCATGCCACCACTACTA

FIG. 18a.

TGTAATTCCCAGCAATTTGGGGAGCCCAAGGCGGGCAGATTCATGAGTTCGGGAAGAT
TCGAGACCNTTCCTGGCTAAACACGGGGGAAACCCCCNTTTTTACTAAAAAATACCAAAA
AATTAACCTGGGCGTGGTGGCGGGCCCCAGCTANTCCGGAGGCTGAGGCAGGAGAAT
GGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCCCGCTACTGCACTCCA
GCCTGGGCAATAGAGGGAGACTCCGTCTCAAAAAAAAAAAAAAAATAAATAATAATAAAA
AAAATAACAATAATAATACTAATAATTGCTTGATATTTTACAAAAGCAAAAGGAAAAGAAG
ACTAGGCAAGAAAAAAAAAACCTCCTTAGATGGTAGAACTCAGGTTTAAAATTAAAACTT
ATTCTGGTGTCAGSCTAGTTGTATATTTTGACCTCTTTAAATGCTCTGAACTATGATATGG
*AGTAACAGCGATG***CTGCTGCTGCTGCTGCTGCTGCTG**ATGGTGGTGGTGTTTTA
ATATCGAATAAAAGTTGTGGAAACTAAATTTCATTTCTGCCAATTAACTAAGATT
GCAAAGTTAAACATCT

FIG. 18b.

5'-TTTGCAATCTTAGTTAATTGGC
5'-GAACTATGATATGGAGTAACAGCG

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION** see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
SCB/48464/001		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/EP 98/ 08543	17/12/1998	18/12/1997
Applicant		
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

   a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

   [ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

   b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

   [ ] contained in the international application in written form.

   [ ] filed together with the international application in computer readable form.

   [X] furnished subsequently to this Authority in written form.

   [X] furnished subsequently to this Authority in computer readble form.

   [X] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

   [X] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).

3. [ ] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

   [X] the text is approved as submitted by the applicant.

   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

   [X] the text is approved as submitted by the applicant.

   [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. _____

   [ ] as suggested by the applicant.

   [ ] because the applicant failed to suggest a figure.

   [ ] because this figure better characterizes the invention.

   [ ] None of the figures.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 98/08543

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/85 C07K14/47 C07K16/18 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 37043 A (UNIV CALIFORNIA ;FRIEMER NELSON B (US); LEON PEDRO (CR); REUS VICT) 9 October 1997 (1997-10-09) the whole document ---	1-3,6, 14, 16-20,25
X	EWALD H ET AL: "SUSCEPTIBILITY LOCI FOR BIPOLAR AFFECTIVE DISORDER ON CHROMOSOME 18? A REVIEW AND A STUDY OF DANISH FAMILIES" PSYCHIATRIC GENETICS, vol. 7, 1997, pages 1-12, XP002911589 ISSN: 0955-8829 the whole document ---	1,7-9

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 August 1999	30/08/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 98/08543

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	BRESCHEL T.S. ET AL.,: "A novel heritable expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1" HUMAN MOLECULAR GENETICS, vol. 6, no. 11, - 11 October 1997 (1997-10-11) pages 1855-1863, XP002112411	1,7-10
Y	see whole doc. esp. M&M	11-13
Y	WO 97 17445 A (CENTRE NAT RECH SCIENT ;INST NAT SANTE RECH MED (FR); TORA LAZSLO) 15 May 1997 (1997-05-15) cited in the application see whole doc. esp. claims	11-13
A,P	TURECKI G ET AL: "eVIDENCEFOR A ROLE OF PHOSPHOLIPASE C-GAMMA.1 IN THE PATHOGENESIS OF BIPOLAR DISORDER" MOLECULAR PSYCHIATRY, vol. 3, no. 6, 1 January 1998 (1998-01-01), pages 534-538, XP002091617 ISSN: 1359-4184 the whole document	
A	O'DONOVAN M.C. ET AL.,: "Expanded CAG repeats in schizophrenia and bipolar disorder" NATURE GENETICS, vol. 10, - August 1995 (1995-08) pages 380-381, XP002112412 cited in the application the whole document	
A	WO 97 11175 A (MEDICAL RES COUNCIL ;BATTERSBY SHARON (GB); FINK GEORGE (GB); GOOD) 27 March 1997 (1997-03-27) see whole doc. esp. claims	
T	VERHEYEN G.R. ET AL.,: "Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder" EUROPEAN JOURNAL OF HUMAN GENETICS, vol. 7, no. 4, - May 1999 (1999-05) pages 427-434, XP002112413 the whole document	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 98/08543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9737043 A	09-10-1997	AU 2423897 A	22-10-1997
		CA 2247996 A	09-10-1997
		AU 4160497 A	06-03-1998
		WO 9807887 A	26-02-1998
WO 9717445 A	15-05-1997	FR 2741088 A	16-05-1997
WO 9711175 A	27-03-1997	AU 7089896 A	09-04-1997
		GB 2321246 A	22-07-1998

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 25 August 1999 (25.08.99)	
International application No. PCT/EP98/08543	**Applicant's or agent's file reference** SCB/48464/001
International filing date (day/month/year) 17 December 1998 (17.12.98)	**Priority date** (day/month/year) 18 December 1997 (18.12.97)
Applicant VAN BROECKHOVEN, Christine et al	

1. The designated Office is hereby notified of its election made:

   [X] in the demand filed with the International Preliminary Examining Authority on:

   13 July 1999 (13.07.99)

   [ ] in a notice effecting later election filed with the International Bureau on:

   _____

2. The election [X] was

   [ ] was not

   made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

A. Karkachi

Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
SCB/48464/001		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/EP 98/ 08543	17/12/1998	18/12/1997

Applicant

VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

[ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

[ ] contained in the international application in written form.

[ ] filed together with the international application in computer readable form.

[X] furnished subsequently to this Authority in written form.

[X] furnished subsequently to this Authority in computer readble form.

[X] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[X] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).

3. [ ] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. _____

[ ] as suggested by the applicant.

[ ] because the applicant failed to suggest a figure.

[ ] because this figure better characterizes the invention.

[ ] None of the figures.

INTERNATIONAL SEARCH REPORT

[redacted] International Application No
[redacted]T/EP 98/08543

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/85 C07K14/47 C07K16/18 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 37043 A (UNIV CALIFORNIA ;FRIEMER NELSON B (US); LEON PEDRO (CR); REUS VICT) 9 October 1997 (1997-10-09) the whole document ---	1-3,6, 14, 16-20,25
X	EWALD H ET AL: "SUSCEPTIBILITY LOCI FOR BIPOLAR AFFECTIVE DISORDER ON CHROMOSOME 18? A REVIEW AND A STUDY OF DANISH FAMILIES" PSYCHIATRIC GENETICS, vol. 7, 1997, pages 1-12, .XP002911589 ISSN: 0955-8829 the whole document --- -/--	1,7-9

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 August 1999	30/08/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	BRESCHEL T.S. ET AL.,: "A novel heritable expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1" HUMAN MOLECULAR GENETICS, vol. 6, no. 11, - 11 October 1997 (1997-10-11) pages 1855-1863, XP002112411	1,7-10
Y	see whole doc. esp. M&M	11-13
Y	WO 97 17445 A (CENTRE NAT RECH SCIENT ;INST NAT SANTE RECH MED (FR); TORA LAZSLO) 15 May 1997 (1997-05-15) cited in the application see whole doc. esp. claims	11-13
A,P	TURECKI G ET AL: "eVIDENCEFOR A ROLE OF PHOSPHOLIPASE C-GAMMA.1 IN THE PATHOGENESIS OF BIPOLAR DISORDER" MOLECULAR PSYCHIATRY, vol. 3, no. 6, 1 January 1998 (1998-01-01), pages 534-538, XP002091617 ISSN: 1359-4184 the whole document	
A	O'DONOVAN M.C. ET AL.,: "Expanded CAG repeats in schizophrenia and bipolar disorder" NATURE GENETICS, vol. 10, - August 1995 (1995-08) pages 380-381, XP002112412 cited in the application the whole document	
A	WO 97 11175 A (MEDICAL RES COUNCIL ;BATTERSBY SHARON (GB); FINK GEORGE (GB); GOOD) 27 March 1997 (1997-03-27) see whole doc. esp. claims	
T	VERHEYEN G.R. ET AL.,: "Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder" EUROPEAN JOURNAL OF HUMAN GENETICS, vol. 7, no. 4, - May 1999 (1999-05) pages 427-434, XP002112413 the whole document	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
/EP 98/08543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9737043 A	09-10-1997	AU 2423897 A	22-10-1997
		CA 2247996 A	09-10-1997
		AU 4160497 A	06-03-1998
		WO 9807887 A	26-02-1998
WO 9717445 A	15-05-1997	FR 2741088 A	16-05-1997
WO 9711175 A	27-03-1997	AU 7089896 A	09-04-1997
		GB 2321246 A	22-07-1998

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
SCB/48464/001		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/EP98/08543	17/12/1998	18/12/1997
International Patent Classification (IPC) or national classification and IPC C12N15/85		
Applicant VLAAMS INTERUNIVERSITAIR INST.VOOR BIOTECHN.et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☒ Certain documents cited
   - VII ☒ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
13/07/1999	13.03.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer SCHEFFZYK, I Telephone No. +49 89 2399 8602

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP98/08543

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-67 as originally filed

Claims, No.:

1-47 as originally filed

Drawings, sheets:

1/7-7/7 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	4, 5,11-13,15, 21-24, 26-32, 36-46
	No:	Claims	1-3, 6-10, 14, 16-20, 25, 33-35,47
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-47
Industrial applicability (IA)	Yes:	Claims	1-47
	No:	Claims	

2. Citations and explanations

see separate sheet

VI. Certain documents cited

1. Certain published documents (Rule 70.10)

and / or

2. Non-written disclosures (Rule 70.9)

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

SECTION V----------

Novelty:

The subject-matter of claims 1-3, 6, 14, 16-20 and 25 is anticipated by the teaching of WO 97/37043 (1) (see e.g. the abstract and the examples). Moreover, Ewald et al. (2) and Breschel et al. (3) also are relevant with respect to novelty of present claims 1 and 7-10, 17, 18 and 19. Moreover, due to the term "related disorder" written in claim 33 the disclosure of WO 97/17445 (4) is considered to be novelty destroying for the subject-matter of claims 33-35. In addition, novelty of claim 47 also is questionable since it cannot be ruled out that readily available compounds are covered by the scope of said claim.
Correspondingly, claims 1-3, 6-10, 14, 16-20, 25, 33-35 and 47 do not meet the requirements of Art. 33(2) PCT.

Claims 4, 5, 7-13, 15, 21-24, 26-32, 36-46 are deemed novel since the embodiments thereof are not disclosed in the available prior art.

Inventive step:

Considering that it is already taught in the prior art that the region specified in claim 1 of the human chromosome 18q is associated with mood disorders (see above) and considering that it is also well-known in the art that such disorders are related with trinucleotide repeats (see e.g. 2) and O'Donovan et al., Nature Genetics, vol. 10, August 1995 (5) the subject-matter of present claims cannot be considered to be inventive but merely as obvious embodiment which arises out of the knowledge of the prior art. Thus, present claims do not meet the requirements of Art. 33(3) PCT.

SECTION VI----------

Turecki G. et al., Molecular Psychiatry, vol.3, no. 6, 01.01.98, pp. 534-538

Verheyen G.R. et al., European Journal of Human Genetics, vol. 7, no. 4, May 1999, pp. 427-434

SECTION VII----------

Claims 14 and 16 do not comply with the requirements of Rule 6.2(a) PCT (...contig described herein) .

SECTION VIII----------

1). Claim 47 is not supported by the specification since the application as filed fails to specify compounds which are covered by the scope of said claim (Art. 6 PCT).

2). In so far as the clones specified in claim 4 have not been deposited the question arises whether present application meets the requirements of Art. 5 PCT with respect to said claim.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
SCB/48464/001		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/EP98/08543	17/12/1998	18/12/1997
International Patent Classification (IPC) or national classification and IPC C12N15/85		
Applicant VLAAMS INTERUNIVERSITAIR INST.VOOR BIOTECHN.et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☒ Certain documents cited
   - VII ☒ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
13/07/1999	13.03.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer SCHEFFZYK, I Telephone No. +49 89 2399 8602

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP98/08543

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-67 as originally filed

Claims, No.:

1-47 as originally filed

Drawings, sheets:

1/7-7/7 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	4, 5,11-13,15, 21-24, 26-32, 36-46
	No:	Claims	1-3, 6-10, 14, 16-20, 25, 33-35,47
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-47
Industrial applicability (IA)	Yes:	Claims	1-47
	No:	Claims	

2. Citations and explanations

see separate sheet

VI. Certain documents cited

1. Certain published documents (Rule 70.10)

and / or

2. Non-written disclosures (Rule 70.9)

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

SECTION V----------

Novelty:

The subject-matter of claims 1-3, 6, 14, 16-20 and 25 is anticipated by the teaching of WO 97/37043 (1) (see e.g. the abstract and the examples). Moreover, Ewald et al. (2) and Breschel et al. (3) also are relevant with respect to novelty of present claims 1 and 7-10, 17, 18 and 19. Moreover, due to the term "related disorder" written in claim 33 the disclosure of WO 97/17445 (4) is considered to be novelty destroying for the subject-matter of claims 33-35. In addition, novelty of claim 47 also is questionable since it cannot be ruled out that readily available compounds are covered by the scope of said claim.
Correspondingly, claims 1-3, 6-10, 14, 16-20, 25, 33-35 and 47 do not meet the requirements of Art. 33(2) PCT.

Claims 4, 5, 7-13, 15, 21-24, 26-32, 36-46 are deemed novel since the embodiments thereof are not disclosed in the available prior art.

Inventive step:

Considering that it is already taught in the prior art that the region specified in claim 1 of the human chromosome 18q is associated with mood disorders (see above) and considering that it is also well-known in the art that such disorders are related with trinucleotide repeats (see e.g. 2) and O'Donovan et al., Nature Genetics, vol. 10, August 1995 (5) the subject-matter of present claims cannot be considered to be inventive but merely as obvious embodiment which arises out of the knowledge of the prior art. Thus, present claims do not meet the requirements of Art. 33(3) PCT.

SECTION VI----------

Turecki G. et al., Molecular Psychiatry, vol.3, no. 6, 01.01.98, pp. 534-538

Verheyen G.R. et al., European Journal of Human Genetics, vol. 7, no. 4, May 1999, pp. 427-434

SECTION VII----------

Claims 14 and 16 do not comply with the requirements of Rule 6.2(a) PCT (...contig described herein) .

SECTION VIII----------

1). Claim 47 is not supported by the specification since the application as filed fails to specify compounds which are covered by the scope of said claim (Art. 6 PCT).

2). In so far as the clones specified in claim 4 have not been deposited the question arises whether present application meets the requirements of Art. 5 PCT with respect to said claim.

PCT

REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only

International Application No.

International Filing Date

Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference *(if desired) (12 characters maximum)* SCB/48464/001

Box No. I TITLE OF INVENTION

MOOD DISORDER GENE

Box No. II APPLICANT

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR
BIOTECHNOLOGIE
RIJVISSCHESTRAAT 118 BUS 1
B-9052 ZWIJNAARDE
BELGIUM

☐ This person is also inventor.

Telephone No.

Facsimile No.

Teleprinter No.

State *(that is, country)* of nationality: BELGIAN

State *(that is, country)* of residence: BELGIUM

This person is applicant for the purposes of: ☐ all designated States ☒ all designated States except the United States of America ☐ the United States of America only ☐ the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

VAN BROECKHOVEN; Christine
KONING ALBERTLEI 15
B - 2650 EDEGEM
BELGIUM

This person is:

☐ applicant only

☒ applicant and inventor

☐ inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: BELGIAN

State *(that is, country)* of residence: BELGIUM

This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America ☒ the United States of America only ☐ the States indicated in the Supplemental Box

☒ Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: ☒ agent ☐ common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

BALDOCK; Sharon Claire
BOULT WADE TENNANT
27 FURNIVAL STREET
LONDON EC4A 1PQ

Telephone No. 0171-404-5921

Facsimile No. 0171-405-1916

Teleprinter No. 267271 BOULT G

☐ **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

See Notes to the request form

Continuation of Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

RAEYMAEKERS; Peter
KARDINAAL CARDIJNLAAN 104
B - 2547 LINT
BELGIUM

This person is:
- [ ] applicant only
- [X] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: BELGIAN

State *(that is, country)* of residence: BELGIUM

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [X] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

DEL-FAVERO; Jurgen
COORBEEKSESTEENWEG 149
B - 330 TIENEN
BELGIUM

This person is:
- [ ] applicant only
- [X] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: BELGIAN

State *(that is, country)* of residence: BELGIUM

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [X] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [ ] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [ ] the United States of America only
- [ ] the States indicated in the Supplemental Box

- [ ] Further applicants and/or (further) inventors are indicated on another continuation sheet.

Box No.V DESIGNATION OF STATES

The following designations are hereby made under Rule 4.9(a) *(mark the applicable check-boxes; at least one must be marked)*:

Regional Patent

- ☒ AP **ARIPO Patent:** GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, SD Sudan, SZ Swaziland, UG Uganda, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT
- ☒ EA **Eurasian Patent:** AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian Patent Convention and of the PCT
- ☒ EP **European Patent:** AT Austria, BE Belgium, CH and LI Switzerland and Liechtenstein, CY Cyprus, DE Germany, DK Denmark, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, and any other State which is a Contracting State of the European Patent Convention and of the PCT
- ☒ OA **OAPI Patent:** BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* .....

National Patent *(if other kind of protection or treatment desired, specify on dotted line)*:

- ☒ AL Albania
- ☒ AM Armenia
- ☒ AT Austria
- ☒ AU Australia
- ☒ AZ Azerbaijan
- ☒ BA Bosnia and Herzegovina
- ☒ BB Barbados
- ☒ BG Bulgaria
- ☒ BR Brazil
- ☒ BY Belarus
- ☒ CA Canada
- ☒ CH and LI Switzerland and Liechtenstein
- ☒ CN China
- ☒ CU Cuba
- ☒ CZ Czech Republic
- ☒ DE Germany
- ☒ DK Denmark
- ☒ EE Estonia
- ☒ ES Spain
- ☒ FI Finland
- ☒ GB United Kingdom
- ☒ GE Georgia
- ☒ GH Ghana
- ☒ GM Gambia
- ☒ GW Guinea-Bissau
- ☒ HR Croatia
- ☒ HU Hungary
- ☒ ID Indonesia
- ☒ IL Israel
- ☒ IS Iceland
- ☒ JP Japan
- ☒ KE Kenya
- ☒ KG Kyrgyzstan
- ☒ KP Democratic People's Republic of Korea
- ☒ KR Republic of Korea
- ☒ KZ Kazakhstan
- ☒ LC Saint Lucia
- ☒ LK Sri Lanka
- ☒ LR Liberia
- ☒ LS Lesotho
- ☒ LT Lithuania
- ☒ LU Luxembourg
- ☒ LV Latvia
- ☒ MD Republic of Moldova
- ☒ MG Madagascar
- ☒ MK The former Yugoslav Republic of Macedonia
- ☒ MN Mongolia
- ☒ MW Malawi
- ☒ MX Mexico
- ☒ NO Norway
- ☒ NZ New Zealand
- ☒ PL Poland
- ☒ PT Portugal
- ☒ RO Romania
- ☒ RU Russian Federation
- ☒ SD Sudan
- ☒ SE Sweden
- ☒ SG Singapore
- ☒ SI Slovenia
- ☒ SK Slovakia
- ☒ SL Sierra Leone
- ☒ TJ Tajikistan
- ☒ TM Turkmenistan
- ☒ TR Turkey
- ☒ TT Trinidad and Tobago
- ☒ UA Ukraine
- ☒ UG Uganda
- ☒ US United States of America
- ☒ UZ Uzbekistan
- ☒ VN Viet Nam
- ☒ YU Yugoslavia
- ☒ ZW Zimbabwe

Check-boxes reserved for designating States (for the purposes of a national patent) which have become party to the PCT after issuance of this sheet:

- ☒ GD Grenada
- ☒ IN India

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. *(Confirmation of a designation consists of the filing of a notice specifying that designation and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.)*

Box No. VI PRIORITY CLAIM ☐ Further priority claims are indicated in the Supplemental Box.

Filing date of earlier application (day/month/year)	Number of earlier application	Where earlier application is: national application: country	regional application:* regional Office	international application: receiving Office
item (1) 18TH DECEMBER 1997	9726804.9	UNITED KINGDOM		
item (2)				
item (3)				

☐ The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of the present international application is the receiving Office)* identified above as item(s): _____

** Where the earlier application is an ARIPO application, it is mandatory to indicate in the Supplemental Box at least one country party to the Paris Convention for the Protection of Industrial Property for which that earlier application was filed (Rule 4.10(b)(ii)). See Supplemental Box.*

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used)*

ISA / EPO

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority)*

Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII CHECK LIST; LANGUAGE OF FILING

This international application contains the following **number of sheets**:

request	:	5
description (excluding sequence listing part)	:	67
claims	:	11
abstract	:	1
drawings	:	7
sequence listing part of description	:	91
Total number of sheets	:	

This international application is **accompanied** by the item(s) marked below:

1. ☒ fee calculation sheet
2. ☐ separate signed power of attorney
3. ☐ copy of general power of attorney; reference number, if any:
4. ☐ statement explaining lack of signature
5. ☒ priority document(s) identified in Box No. VI as item(s):
6. ☐ translation of international application into (language):
7. ☐ separate indications concerning deposited microorganism or other biological material
8. ☐ nucleotide and/or amino acid sequence listing in computer readable form
9. ☒ other *(specify)*: Alternative Representatives

Figure of the drawings which should accompany the abstract: None

Language of filing of the international application: ENGLISH

Box No. IX SIGNATURE OF APPLICANT OR AGENT

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

BALDOCK; Sharon Claire
BOULT WADE TENNANT 16th December 1998

For receiving Office use only

1. Date of actual receipt of the purported international application:	2. Drawings: ☐ received: ☐ not received:
3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:	
4. Date of timely receipt of the required corrections under PCT Article 11(2):	
5. International Searching Authority (if two or more are competent): ISA /	6. ☐ Transmittal of search copy delayed until search fee is paid.

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

ADDITIONAL SHEET

(Euro-PCT Application)

2. Additional Representatives:-

MAYES; Stuart David

BAYLISS; Geoffrey Cyril

ALEXANDER; Thomas Bruce

BUCKS; Teresa Anne

PLUCKROSE; Anthony William

McLEISH; Nicholas Alistair Maxwell

MERRIFIELD; Sarah Elizabeth

BAVERSTOCK; Michael George Douglas

ALLARD; Susan Joyce

CROSS; Rupert Edward Blount

RICKARD; David John

HAYES; Adrian Chetwynd

All of
BOULT WADE TENNANT
27 FURNIVAL STREET
LONDON
EC4A 1PQ
UNITED KINGDOM

Signed_____
BALDOCK; Sharon Claire

Our Ref: SCB/48464/000

This sheet is not part of and does not count as a sheet of the international application.

PCT

FEE CALCULATION SHEET

Annex to the Request

For receiving Office use only

International application No.

Date stamp of the receiving Office

Applicant's or agent's file reference: SCB/48464/000

Applicant: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE

CALCULATION OF PRESCRIBED FEES

1. TRANSMITTAL FEE . . . . . 200 [T]
2. SEARCH FEE . . . . . 2200 [S]

   International search to be carried out by _____

   (If two or more International Searching Authorities are competent in relation to the international application, indicate the name of the Authority which is chosen to carry out the international search.)

3. INTERNATIONAL FEE

   Basic Fee

   The international application contains 91 sheets.

   first 30 sheets . . . . . 800 [b1]

   61 (remaining sheets) x 19 (additional amount) = 1159 [b2]

   Add amounts entered at b1 and b2 and enter total at B . . . . 1959 [B]

   Designation Fees

   The international application contains 88 designations.

   11 (number of designation fees payable *(maximum 11)*) x 184 (amount of designation fee) = 2024 [D]

   Add amounts entered at B and D and enter total at I . . . . . 3983 [I]

   (Applicants from certain States are entitled to a reduction of 75% of the international fee. Where the applicant is (or all applicants are) so entitled, the total to be entered at I is 25% of the sum of the amounts entered at B and D.)

4. FEE FOR PRIORITY DOCUMENT *(if applicable)* . . . . . [P]

5. TOTAL FEES PAYABLE . . . . . 6383 TOTAL

   Add amounts entered at T, S, I and P, and enter total in the TOTAL box

☐ The designation fees are not paid at this time.

MODE OF PAYMENT

- ☒ authorization to charge deposit account (see below)
- ☐ cheque
- ☐ postal money order
- ☐ bank draft
- ☐ cash
- ☐ revenue stamps
- ☐ coupons
- ☐ other *(specify)*:

DEPOSIT ACCOUNT AUTHORIZATION *(this mode of payment may not be available at all receiving Offices)*

The RO/ _____

- ☒ is hereby authorized to charge the total fees indicated above to my deposit account.
- ☒ is hereby authorized to charge any deficiency or credit any overpayment in the total fees indicated above to my deposit account.
- ☐ is hereby authorized to charge the fee for preparation and transmittal of the priority document to the International Bureau of WIPO to my deposit account.

2805.0034	16th December 1998	BALDOCK; Sharon Claire
Deposit Account No.	Date *(day/month/year)*	Signature

See Notes to the fee calculation sheet

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/85, C07K 14/47, 16/18, C12Q 1/68

A3

(11) International Publication Number: WO 99/32643

(43) International Publication Date: 1 July 1999 (01.07.99)

(21) International Application Number: PCT/EP98/08543

(22) International Filing Date: 17 December 1998 (17.12.98)

(30) Priority Data:
9726804.9 18 December 1997 (18.12.97) GB

(71) Applicant *(for all designated States except US)*: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE [BE/BE]; Rijvisschestraat 118 Bus 1, B–9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VAN BROECKHOVEN, Christine [BE/BE]; Koning Albertlei 15, B–2650 Edegem (BE). RAEYMAEKERS, Peter [BE/BE]; Kardinaal Cardijnlaan 104, B–2547 Lint (BE). DEL–FAVERO, Jurgen [BE/BE]; Oorbeeksesteenweg 149, B–3300 Tienen (BE).

(74) Agents: BALDOCK, Sharon, Claire et al.; Boult Wade Tennant, 27 Furnival Street, London EC4A 1PQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
21 October 1999 (21.10.99)

(54) Title: MOOD DISORDER GENE

(57) Abstract

The present invention comprises the use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders. The invention also provides methods for determining the susceptibility of an individual to mood disorders or related disorders, comprising analysing a DNA sample for the presence of a trinucleotide repeat expansion in the above region. Polynucleotide sequences useful for detecting the presence of such trinucleotide repeat expansions are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 98/08543

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/85 C07K14/47 C07K16/18 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 37043 A (UNIV CALIFORNIA ;FRIEMER NELSON B (US); LEON PEDRO (CR); REUS VICT) 9 October 1997 (1997-10-09) the whole document ---	1-3,6, 14, 16-20,25
X	EWALD H ET AL: "SUSCEPTIBILITY LOCI FOR BIPOLAR AFFECTIVE DISORDER ON CHROMOSOME 18? A REVIEW AND A STUDY OF DANISH FAMILIES" PSYCHIATRIC GENETICS, vol. 7, 1997, pages 1-12, XP002911589 ISSN: 0955-8829 the whole document ---	1,7-9

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 August 1999	30/08/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 98/08543

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	BRESCHEL T.S. ET AL.,: "A novel heritable expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1" HUMAN MOLECULAR GENETICS, vol. 6, no. 11, - 11 October 1997 (1997-10-11) pages 1855-1863, XP002112411	1,7-10
Y	see whole doc. esp. M&M	11-13
Y	WO 97 17445 A (CENTRE NAT RECH SCIENT ;INST NAT SANTE RECH MED (FR); TORA LAZSLO) 15 May 1997 (1997-05-15) cited in the application see whole doc. esp. claims	11-13
A,P	TURECKI G ET AL: "eVIDENCEFOR A ROLE OF PHOSPHOLIPASE C-GAMMA.1 IN THE PATHOGENESIS OF BIPOLAR DISORDER" MOLECULAR PSYCHIATRY, vol. 3, no. 6, 1 January 1998 (1998-01-01), pages 534-538, XP002091617 ISSN: 1359-4184 the whole document	
A	O'DONOVAN M.C. ET AL.,: "Expanded CAG repeats in schizophrenia and bipolar disorder" NATURE GENETICS, vol. 10, - August 1995 (1995-08) pages 380-381, XP002112412 cited in the application the whole document	
A	WO 97 11175 A (MEDICAL RES COUNCIL ;BATTERSBY SHARON (GB); FINK GEORGE (GB); GOOD) 27 March 1997 (1997-03-27) see whole doc. esp. claims	
T	VERHEYEN G.R. ET AL.,: "Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder" EUROPEAN JOURNAL OF HUMAN GENETICS, vol. 7, no. 4, - May 1999 (1999-05) pages 427-434, XP002112413 the whole document	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 98/08543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9737043 A	09-10-1997	AU 2423897 A	22-10-1997
		CA 2247996 A	09-10-1997
		AU 4160497 A	06-03-1998
		WO 9807887 A	26-02-1998
WO 9717445 A	15-05-1997	FR 2741088 A	16-05-1997
WO 9711175 A	27-03-1997	AU 7089896 A	09-04-1997
		GB 2321246 A	22-07-1998

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/85, C07K 14/47, 16/18, C12Q 1/68

A2

(11) International Publication Number: WO 99/32643

(43) International Publication Date: 1 July 1999 (01.07.99)

(21) International Application Number: PCT/EP98/08543

(22) International Filing Date: 17 December 1998 (17.12.98)

(30) Priority Data:
9726804.9 18 December 1997 (18.12.97) GB

(71) Applicant *(for all designated States except US)*: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE [BE/BE]; Rijvisschestraat 118 Bus 1, B–9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VAN BROECKHOVEN, Christine [BE/BE]; Koning Albertlei 15, B–2650 Edegem (BE). RAEYMAEKERS, Peter [BE/BE]; Kardinaal Cardijnlaan 104, B–2547 Lint (BE). DEL–FAVERO, Jurgen [BE/BE]; Oorbeeksesteenweg 149, B–3300 Tienen (BE).

(74) Agents: BALDOCK, Sharon, Claire et al.; Boult Wade Tennant, 27 Furnival Street, London EC4A 1PQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MOOD DISORDER GENE

(57) Abstract

The present invention comprises the use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders. The invention also provides methods for determining the susceptibility of an individual to mood disorders or related disorders, comprising analysing a DNA sample for the presence of a trinucleotide repeat expansion in the above region. Polynucleotide sequences useful for detecting the presence of such trinucleotide repeat expansions are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

09/581500

MOOD DISORDER GENE

The invention is concerned with the determination of genetic factors associated with psychiatric health with particular reference to a human gene or genes which contributes to or is responsible for the manifestation of a mood disorder or a related disorder in affected individuals. In particular, although not exclusively, the invention provides a method of identifying and characterising such a gene or genes from human chromosome 18, as well as genes so identified and their expression products. The invention is also concerned with methods of determining the genetic susceptibility of an individual to a mood disorder or related disorder. By mood disorders or related disorders is meant the following disorders as defined in the Diagnostic and Statistical Manual of Mental Disorders, version 4 (DSM-IV) taxonomy (DSM-IV codes in parenthesis):- mood disorders (296.XX, 300.4, 311, 301.13, 295.70), schizophrenia and related disorders (295.XX, 297.1,298.8, 297.3, 298.9), anxiety disorders (300.XX, 309.81,308.3), adjustment disorders (309.XX) and personality disorders (codes 301.XX).

The methods of the invention are particularly exemplified in relation to genetic factors associated with a family of mood disorders known as Bipolar (BP) spectrum disorders.

Bipolar disorder (BP) is a severe psychiatric condition that is characterized by disturbances in mood, ranging from an extreme state of elation (mania) to a severe state of dysphoria (depression). Two types of bipolar illness have been described: type I BP illness (BPI) is characterized by major depressive episodes alternated with phases of mania, and type II

BP illness (BPII), characterized by major depressive episodes alternating with phases of hypomania. Relatives of BP probands have an increased risk for BP, unipolar disorder (patients only experiencing depressive episodes; UP), cyclothymia (minor depression and hypomania episodes; CY) as well as for schizoaffective disorders of the manic (SAm) and depressive (SAd) type. Based on these observations BP, CY, UP and SA are classified as BP spectrum disorders. The involvement of genetic factors in the etiology of BP spectrum disorders was suggested by family, twin and adoption studies (Tsuang and Faraone (1990), The Genetics of Mood Disorders, Baltimore, The John Hopkins University Press). However, the exact pattern of transmission is unknown. In some studies, complex segregation analysis supports the existence of a single major locus for BP (Spence *et al.* (1995), Am J. Med. Genet (Neuropsych. Genet.) 60 pp 370-376). Other researchers propose a liability-threshold-model, in which the liability to develop the disorder results from the additive combination of multiple genetic and environmental effects (McGuffin *et al.* (1994), Affective Disorders; Seminars in Psychiatric Genetics Gaskell, London pp 110-127).

Due to the complex mode of inheritance, parametric and nonparametric linkage strategies are applied in families in which BP disorder appears to be transmitted in a Mendelian fashion. Early linkage findings on chromosomes 11p15 (Egeland *et al.* (1987), Nature **325** pp 783-787) and Xq27-q28 (Mendlewicz *et al.* (1987) The Lancet **1** pp 1230 -1232; Baron *et al.* (1987) Nature **326** pp 289-292) have been controversial and could initially not be replicated (Kelsoe *et al.* (1989) Nature **242** pp 238-243; Baron *et al.* (1993) Nature Genet **3** pp 49-55). With the development of a

human genetic map saturated with highly polymorphic markers and the continuous development of data analysis techniques, numerous new linkage searches were started. In several studies, evidence or suggestive evidence for linkage to particular regions on chromosomes 4, 12, 18, 21 and X was found (Blackwood *et al.* (1996) Nature Genetics **12** pp 427-430, Craddock *et al.* (1994) Brit J. Psychiatry **164** pp 355-358, Berrettini *et al.* (1994), Proc Natl Acad Sci USA **91** pp 5918-5921, Straub *et al.* (1994) Nature Genetics **8** pp 291-296 and Pekkarinen *et al.* (1995) Genome Research **5** pp 105-115). In order to test the validity of the reported linkage results, these findings have to be replicated in other, independent studies.

Recently, linkage of bipolar disorder to the pericentromeric region on chromosome 18 was reported (Berrettini *et al.* 1994). Also a ring chromosome 18 with break-points and deleted regions at 18pter-p11 and 18q23-qter was reported in three unrelated patients with BP illness or related syndromes (Craddock *et al.* 1994). The chromosome 18p linkage was replicated by Stine *et al.* (1995) Am J Hum Genet **57** pp 1384-1394, who also reported suggestive evidence for a locus on 18q21.2-q21.32 in the same study. Interestingly, Stine *et al.* observed a parent-of-origin effect: the evidence of linkage was the strongest in the paternal pedigrees, in which the proband's father or one of the proband's father's sibs is affected.

In an independent replication study, the present inventors tested linkage with chromosome 18 markers in 10 Belgian families with a bipolar proband. To localize causative genes the linkage analysis or likelihood method was used in these families. This

method studies within a family the segregation of a defined disease phenotype with that of polymorphic genetic markers distributed in the human genome. The likelihood ratio of observing cosegregation of the disease and a genetic marker under linkage versus no linkage is calculated and the log of this ratio or the log of the odds is the LOD score statistic z. A LOD score of 3 (or likelihood ratio of 1000 or greater) is taken as significant statistical evidence for linkage. In the inventors' study no evidence for linkage to the pericentromeric regions was found, but in one of the families, MAD31, a Belgian family of a BPII proband, suggestive linkage was found with markers located at 18q21.33-q23 (De bruyn *et al.* (1996) Biol Psychiatry **39** pp 679-688). Multipoint linkage analysis gave the highest LOD score in the interval between STR (Short Tandem Repeats) polymorphisms D18S51 and D18S61, with a maximum multipoint LOD score of +1.34. Simulation studies indicated that this LOD score is within the range of what can be expected for a linked marker given the information available in the family. Likewise, an affected sib-pair analysis also rejected the null-hypothesis of nonlinkage for several of the markers tested. Two other groups also found evidence for linkage of bipolar disorder to 18q (Freimer *et al.* (1996) Nature Genetics **12** pp 436-441, Coon *et al.* (1996) Biol Psychiatry **39** pp 689 to 696). Although the candidate regions in the different studies do not entirely overlap, they all suggest the presence of a susceptibility locus at 18q21-q23.

The inventors have now carried out further investigations into the 18q chromosomal region in family MAD31. By analysis of cosegregation of bipolar disease in MAD31 with twelve STR polymorphic markers previously located between the aforementioned markers

D18S51 and D18S61 and subsequent LOD score analysis as described above, the inventors have further refined the candidate region of chromosome 18 in which a gene associated with mood disorders such as bipolar spectrum disorders may be located and have constructed a physical map. The region in question may thus be used to locate, isolate and sequence a gene or genes which influences psychiatric health and mood.

The inventors have also constructed a YAC (yeast artificial chromosome) contig map of the candidate region to determine the relative order of the twelve STR markers mapped by the cosegregational analysis and they have identified seven clones from the YAC library incorporating the candidate region.

A number of procedures can be applied to the identified YAC clones and, where applicable, to the DNA of an individual afflicted with a mood disorder as defined herein, in the process of identifying and characterising the relevant gene or genes. For example, the inventors have used YAC clones spanning the region of interest in chromosome 18 to identify by CAG or CTG fragmentation novel genes that are allegedly involved in the manifestation of mood disorders or related disorders.

Other procedures can also be applied to the said YAC clones to identify candidate genes as discussed below.

Once candidate genes have been identified it is possible to assess the susceptibility of an individual to a mood disorder or related disorder by detecting the presence of a polymorphism associated with a mood disorder or related disorder in such genes.

Accordingly, in a first aspect the present invention comprises the use of an 8.9 cM region of

human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with mood disorders or related disorders as defined above. As will be described below, the present inventors have identified this candidate region of chromosome 18q for such a gene, by analysis of cosegregation of bipolar disease in family MAD31 with 12 STR polymorphic markers previously located between D18S51 and D18S61 and subsequent LOD score analysis.

In a second aspect the invention comprises the use of a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61 for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with mood disorders or related disorders as defined above. D18S60 is close to D18S51 so the particular YAC clones for use are those which have an artificial chromosome spanning the candidate region of human chromosome 18q between polymorphic markers D18S51 and D18S61 as identified by the present inventors in their earlier paper (De bruyn *et al.* (1996)).

Particular YACs covering the candidate region which may be used in accordance with the present invention are 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752-g-8 and 717_d_3, preferred ones being 961_h_9, 766_f_12 and 907_e_1 since these have the minimum tiling path across the candidate region. Suitable YAC clones for use are those having an artificial chromosome spanning the refined candidate region between D18S68 and D18S979.

There are a number of methods which can be applied to the candidate regions of chromosome 18q as defined above, whether or not present in a YAC, to identify a candidate gene or genes associated with mood disorders or related disorders. For example, it has previously been demonstrated that an apparent association exists between the presence of trinucleotide repeat expansions (TRE) in the human genome and the phenomenon of anticipation of mood disorders (Lindblad *et al.* (1995), Neurobiology of Disease **2**: 55-62 and O'Donovan *et al.* (1995), Nature Genetics **10**: 380-381).

Accordingly, in a third aspect the present invention comprises a method of identifying at least one human gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder as defined herein which comprises detecting nucleotide triplet repeats in the region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

An alternative method of identifying said gene or genes comprises fragmenting a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61, for example one or more of the seven aforementioned YAC clones, and detecting any nucleotide triplet repeats in said fragments. Nucleic acid probes comprising at least 5 and preferably at least 10 CTG and/or CAG triplet repeats are a suitable means of detection when appropriately labelled. Trinucleotide repeats may also be determined using the known RED (repeat expansion detection) system (Shalling et al.(1993), Nature Genetics 4 pp 135-139).

In a fourth embodiment the invention comprises a

method of identifying at least one gene, including mutated and polymorphic variants thereof, which is associated with a mood disorder or related disorder and which is present in a YAC clone spanning the region of human chromosome 18q between polymorphic markers D18S60 and D18S61, the method comprising the step of detecting the expression product of a gene incorporating nucleotide triplet repeats by use of an antibody capable of recognising a protein with an amino acid sequence comprising a string of at least 8, but preferably at least 12, continuous glutamine residues. Such a method may be implemented by subcloning YAC DNA, for example from the seven aforementioned YAC clones, into a human DNA expression library. A preferred means of detecting the relevant expression product is by use of a monoclonal antibody, in particular mAB 1C2, the preparation and properties of which are described in International Patent Application Publication No WO 97/17445.

As will be described in detail below, in order to identify candidate genes containing triplet repeats, the inventors have carried out direct CAG or CTG fragmentation of YACs 961_h_9, 766_f_12 and 907_e_1, comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61, and have identified a number of sequences containing CAG or CTG repeats, whose abnormal expansion may be involved in genetic susceptibility to a mood disorder or related disorder.

Accordingly, in a fifth aspect, the invention provides a nucleic acid comprising the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a.

In a further aspect, the invention provides a protein comprising an amino acid sequence encoded by the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a.

In yet a further aspect the invention provides a mutated nucleic acid comprising a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats.

Also provided by the invention is a mutated protein comprising an amino acid sequence encoded by a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats.

It is to be understood that the invention also contemplates nucleotide sequences having at least 75% and preferably at least 80% homology with any of the sequences described above and having functional identity with any of said sequences. The homology is calculated as described by Altschul *et al.* (1997) Nucleic Acids Res. **25**: 3389-3402, Karlin *et al.* (1990) Proc Natl Acad Sci USA **87**: 2264-68 and Karlin *et al.* (1993) Proc Natl Acad Sci USA **90**: 5873-5877. Also contemplated are amino acid sequences which differ from the above described sequences only in conservative amino acid changes. Suitable changes are well known to those skilled in the art.

Knowledge of the sequences described above can be used to design assays to determine the genetic susceptibility of an individual to a mood disorder or

related disorder.

Accordingly, in a further aspect the invention provides a method for determining the susceptibility of an individual to a mood disorder or related disorder which comprises the steps of:

a) obtaining a DNA sample from said individual;

b) providing primers suitable for the amplification of a nucleotide sequence comprised in the sequence shown in any one of Figures 15a, 16a, 17a or 18a said primers flanking the trinucleotide repeats comprised in said sequence;

c) applying said primers to the said DNA sample and carrying out an amplification reaction;

d) carrying out the same amplification reaction on a DNA sample from a control individual; and

e) comparing the results of the amplification reaction for the said individual and for the said control individual;

wherein the presence of an amplified fragment from said individual which is bigger in size from that of said control individual is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.
By control individual is meant an individual who is not affected by a mood disorder or related disorder and does not have a family history of mood disorders or related disorders.

Preferable primers to use in this method are those shown in Figure 15b, 16b, 17b or 18b but other suitable primers may be utilised.

In a further aspect the invention provides a method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises the steps of :

a) obtaining a protein sample from said individual; and

b) detecting the presence of a protein comprising an amino acid sequence encoded by a sequence of nucleotides which differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a, or 18a only in the extent of trinucleotide repeats

wherein the presence of said protein is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

Preferably, the foresaid protein is detected by utilising an antibody that is capable of recognising a string of at least 8 continuous glutamines as, for example, the mAB 1C2 antibody.

The nucleic acids molecules according to the invention may be advantageously included in an expression vector, which may be introduced into a host cell of prokaryotic or eukaryotic origin. Suitable expression vectors include plasmids, which may be used to express foreign DNA in bacterial or eukaryotic host cells, viral vectors, yeast artificial chromosomes or mammalian artificial chromosomes. The vector may be

transfected or transformed into host cells using suitable methods known in the art such as, for example, electroporation, microinjection, infection, lipoinfection and direct uptake. Such methods are described in more detail, for example, by Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", 2nd ed. (1989) and by Ausbel *et al.* "Current Protocols in Molecular Biology", (1994).

Also provided by the invention is a host cell, tissue or organism comprising the expression vector according to the invention. The invention further provides a transgenic host cell, tissue or organism comprising a transgene capable of encoding the proteins of the invention, which may comprise a genomic DNA or a cDNA. The transgene may be present in the trangenic host cell, tissue or organism either stably integrated into the genome or in an extra chromosomal state.

A nucleic acid molecule comprising a nucleotide sequence shown in any one of Figures 15a, 16a, 17a or 18a as well as the protein encoded by it may be therapeutically used in the treatment of mood disorders or related disorders in patients which present a trinucleotide repeat expansion (TRE) in at least one of the foresaid sequences.

Accordingly, in another of its aspects the invention provides the above described nucleic acid molecules and proteins for use as medicaments for the treatment of individuals with a mood disorder or related disorder. Preferably, the nucleic acid or the protein is present in an appropriate carrier or delivery vehicle. As an example, the nucleic acid inserted into a vector, for example a plasmid or a

viral vector, may be transfected into a mammalian cell such as a somatic cell or a mammalian germ line cell, as described above. The cell to be transfected can be present in a biological sample obtained from the patient, for example blood or bone marrow, or can be obtained from cell culture. After transfection the sample may be returned or readministered to a patient according to methods known to those practised in the art, for example, methods as described in Kasid *et al.*, Proc. Natl. Acad. Sci. USA (1990) **87**:473; Rosenberg *et al.* (1990) New Eng. J. Med. **323**: 570 ; Williams *et al.* (1994) Nature **310**: 476; Dick *et al.* (1985) Cell **42**:71; Keller *et al.* (1985) Nature **318**: 149 and Anderson *et al.* (1994) US Patent N. 5,399,346.

There are a number of viral vectors known to those skilled in the art which can be used to introduce the nucleic acid into mammalian cells, for example retroviruses, parvoviruses, coronaviruses, negative strand RNA viruses such as picornaviruses or alphaviruses and double stranded DNA viruses including adenoviruses, herpesviruses such as Herpes Simplex virus types 1 and 2, Epstein-Barr virus or cytomegalovirus and poxviruses such as vaccinia fowlpox or canarypox. Other viruses include, for example, Norwalk viruses, togaviruses, flaviviruses, reoviruses, papovaviruses, hepadnaviruses and hepatitis viruses.

A preferred method to introduce nucleic acid that encodes the desired protein into cells is through the use of engineered viral vectors. These vectors provide a means to introduce nucleic acids into cycling and quiescent cells and have been modified to reduce cytotoxicity and to improve genetic stability. The preparation and use of engineered Herpes simplex virus type 1 (D.M. Krisky, *et al.* (1997) Gene Therapy

4(10): 1120-1125), adenoviral (A. Amalfitanl, *et al.*(1998) Journal of Virology **72**(2):926-933), attenuated lentiviral (R. Zufferey, *et al.*, Nature Biotechnology (1997) **15**(9)871-875) and adenoviral/retroviral chimeric (M. Feng, et al, Nature Biotechnology (1997) **15**(9):866-870) vectors are known to the skilled artisan.

The protein may be administered using methods known in the art. For example, the mode of administration is preferably at the location of the target cells. The administration can be by injection. Other modes of administration (parenteral, mucosal, systemic, implant, intraperitoneal, etc.) are generally known in the art. The agents can, preferably, be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringer's solution and isotonic sodium chloride solution.

In yet another of its aspects the invention provides assay methods for identifying compounds that are able to enhance or inhibit the expression of the proteins of the invention. These assays can be conducted, for example, by transfecting a nucleic acid of the invention into host cells and then comparing the levels of mRNA transcript or the levels of protein expressed from said nucleic acids in the presence or absence of the compound.
Different methods, well known to those skilled in the art can be employed in order to measure transcription or expression levels.
Alternatively, it is possible to identify compounds that modulate transcription by using a reporter gene assay of the type well known in the art. In such an assay a reporter plasmid is constructed in which the

promoter of a gene, whose levels of transcription are to be monitored, is positioned upstream of a gene capable of expressing a reporter molecule. The reporter molecule is a molecule whose level of expression can be easily detected and may be either the transcript of the reporter gene or a protein with characteristics that allow it to be detected. For example, the molecule may be a fluorescent protein such as green fluorescent protein (GFP).

Compound assays may be conducted by introducing the reporter plasmid described above into an appropriate host cell and then measuring the amount of reporter molecule expressed in the presence or absence of the compound to be tested.

The invention also relates to compounds identified by the above mentioned methods.

Further embodiments of the present invention relate to methods of identifying the relevant gene or genes which involve the sub-cloning of YAC DNA as defined above into vectors such as BAC (bacterial artificial chromosome) or PAC (P1 or phage artificial chromosome) or cosmid vectors such as exon-trap cosmid vectors. The starting point for such methods is the construction of a contig map of the region of human chromosome 18q between polymorphic markers D18S60 and D18S61. To this end the present inventors have sequenced the end regions of the fragment of human DNA in each of the seven aforementioned YAC clones and these sequences are disclosed herein. Following subcloning of YAC DNA into other vectors as described above, probes comprising these end sequences or portions thereof, in particular those sequences shown in Figures 1 to 11 herein, together with any known

sequenced tagged site (STS) in this region, as described in the YAC clone contig shown herein, as can be used to detect overlaps between said subclones and a contig map can be constructed. Also the known sequences in the current YAC contig can be used for the generation of contig map subclones.

One route by which a gene or genes which is associated with a mood disorder or associated disorder can be identified is by use of the known technique of exon trapping.

This is an artificial RNA splicing assay, most often making use in current protocols of a specialized exon-trap cosmid vector. The vector contains an artificial minigene consisting of a segment of the SV40 genome containing an origin of replication and a powerful promoter sequence, two splicing-competent exons separated by an intron which contains a multiple cloning site and an SV40 polyadenylation site.

The YAC DNA is subcloned in the exon-trap vector and the recombinant DNA is transfected into a strain of mammalian cells. Transcription from the SV40 promoter results in an RNA transcript which normally splices to include the two exons of the minigene. If the cloned DNA itself contains a functional exon, it can be spliced to the exons present in the vector's minigene. Using reverse transcriptase a cDNA copy can be made and using specific PCR primers, splicing events involving exons of the insert DNA can be identified. Such a procedure can identify coding regions in the YAC DNA which can be compared to the equivalent regions of DNA from a person afflicted with a mood disorder or related disorder to identify the relevant gene.

Accordingly, in a further aspect the invention

comprises a method of identifying at least one human gene, including mutated variants and polymorphisms thereof, which is associated with a mood disorder or related disorder which comprises the steps of:

(a) transfecting mammalian cells with exon trap cosmid vectors prepared and mapped as described above;

(b) culturing said mammalian cells in an appropriate medium;

(c) isolating RNA transcripts expressed from the SV40 promoter;

(d) preparing cDNA from said RNA transcripts;

(e) identifying splicing events involving exons of the DNA subcloned into said exon trap cosmid vectors to elucidate positions of coding regions in said subcloned DNA;

(f) detecting differences between said coding regions and equivalent regions in the DNA of an individual afflicted with said mood disorder or related disorder; and

(g) identifying said gene or mutated or polymorphic variant thereof which is associated with said mood disorder or related disorders.

As an alternative to exon trapping the YAC DNA may be subcloned into BAC, PAC, cosmid or other vectors and a contig map constructed as described above. There are a variety of known methods available by which the position of relevant genes on the

subcloned DNA can be established as follows:

(a) cDNA selection or capture (also called direct selection and cDNA selection): this method involves the forming of genomic DNA/cDNA heteroduplexes by hybridizing a cloned DNA (e.g. an insert of a YAC DNA), to a complex mixture of cDNAs, such as the inserts of all cDNA clones from a specific (e.g. brain) cDNA library. Related sequences will hybridize and can be enriched in subsequent steps using biotin-streptavidine capturing and PCR (or related techniques);

(b) hybridization to mRNA/cDNA: a genomic clone (e.g. the insert of a specific cosmid) can be hybridized to a Northern blot of mRNA from a panel of culture cell lines or against appropriate (e.g. brain) cDNA libraries. A positive signal can indicate the presence of a gene within the cloned fragment;

(c) CpG island identification: CpG or HTF islands are short (about 1 kb) hypomethylated GC-rich (> 60%) sequences which are often found at the 5' ends of genes. CpG islands often have restriction sites for several rare-cutter restriction enzymes. Clustering of rare-cutter restriction sites is indicative of a CpG island and therefore of a possible gene. CpG islands can be detected by hybridization of a DNA clone to Southern blots of genomic DNA digested with rare-cutting enzymes, or by island-rescue PCR (isolation of CpG islands from YACs by amplifying sequences between islands and neighbouring *Alu*-repeats);

(d) zoo-blotting: hybridizing a DNA clone (e.g.

the insert of a specific cosmid) at reduced stringency against a Southern blot of genomic DNA samples from a variety of animal species. Detection of hybridization signals can suggest conserved sequences, indicating a possible gene.

Accordingly, in a further aspect the invention comprises a method of identifying at least one human gene including mutated and polymorphic variants thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) subcloning the YAC DNA as described above into a cosmid, BAC, PAC or other vector;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other sequenced tagged site (STS) in this region as in the YAC clone contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to detect overlaps amongst the subclones and construct a map thereof;

(c) identifying the position of genes within the subcloned DNA by one or more of CpG island identification, zoo-blotting, hybridization of the subcloned DNA to a cDNA library or a Northern blot of mRNA from a panel of culture cell lines;

(d) detecting differences between said genes and equivalent region of the DNA of an individual afflicted with a mood disorder or related disorder; and

(e) identifying said gene which is associated

with said mood disorders or related disorders.

If the cloned YAC DNA is sequenced, computer analysis can be used to establish the presence of relevant genes. Techniques such as homology searching and exon prediction may be applied.

Once a candidate gene has been isolated in accordance with the methods of the invention more detailed comparisons may be made between the gene from a normal individual and one afflicted with a mood disorder such as a bipolar spectrum disorder. For example, there are two methods, described as "mutation testing", by which a mutation or polymorphism in a DNA sequence can be identified. In the first the DNA sample may be tested for the presence or absence of one specific mutation but this requires knowledge of what the mutation might be. In the second a sample of DNA is screened for any deviation from a control (normal) DNA. This latter method is more useful for identifying candidate genes where a mutation is not identified in advance.

In addition, the following techniques may be further applied to a gene identified by the above-described methods to identify differences between genes from normal or healthy individuals and those afflicted with a mood disorder or related disorder:

(a) Southern blotting techniques: a clone is hybridized to nylon membranes containing genomic DNA digested with different restriction enzymes of patients and healthy individuals. Large differences between patients and healthy individuals can be visualized using a radioactive labelling protocol;

(b) heteroduplex mobility in polyacrylamide gels: this technique is based on the fact that the mobility of heteroduplexes in non-denaturing polyacrylamide gels is less than the mobility of homoduplexes. It is most effective for fragments under 200 bp;

(c) single-strand conformational polymorphism analysis (SSCP or SSCA): single stranded DNA folds up to form complex structures that are stabilized by weak intramolecular bonds. The electrophoretic mobilities of these structures on non-denaturing polyacrylamide gels depends on their chain lengths and on their conformation;

(d) chemical cleavage of mismatches (CCM): a radiolabelled probe is hybridized to the test DNA, and mismatches detected by a series of chemical reactions that cleave one strand of the DNA at the site of the mismatch. This is a very sensitive method and can be applied to kilobase-length samples;

(e) enzymatic cleavage of mismatches: the assay is similar to CCM, but the cleavage is performed by certain bacteriophage enzymes.

(f) denaturing gradient gel electrophoresis: in this technique, DNA duplexes are forced to migrate through an electrophoretic gel in which there is a gradient of increasing amounts of a denaturant (chemical or temperature). Migration continues until the DNA duplexes reach a position on the gel wherein the strands melt and separate, after which the denatured DNA does not migrate much further. A single base pair difference between a normal and a mutant DNA duplex is sufficient to cause them to migrate to

different positions in the gel;

(g) direct DNA sequencing.

It will be appreciated that with respect to the methods described herein, in the step of detecting differences between coding regions from the YAC and the DNA of an individual afflicted with a mood disorder or related disorder, the said individual may be anybody with the disorder and not necessary a member of family MAD31.

In accordance with further aspects the present invention provides an isolated human gene and variants thereof associated with a mood disorder or related disorder and which is obtainable by any of the above described methods, an isolated human protein encoded by said gene and a cDNA encoding said protein.

In the experimental report which follows reference will be made to the following figures:

FIGURE 1 shows a sequence of nucleotides which is the left arm end-sequence of YAC 766_f_12;

FIGURE 2 shows a sequence of nucleotides which is a right arm end-sequence of YAC 766_f_12;

FIGURE 3 shows a sequence of nucleotides which is a left arm end-sequence of YAC 717_d_3;

FIGURE 4 shows a sequence of nucleotides which is a right arm end-sequence of YAC 717_d_3;

FIGURE 5 shows a sequence of nucleotides which is

a right arm end-sequence of YAC 731_c_7;

FIGURE 6 shows a sequence of nucleotides which is a left arm end-sequence of YAC 752_g_8;

FIGURE 7 shows a sequence of nucleotides which is a left arm end-sequence of YAC 942_c_3;

FIGURE 8 shows a sequence of nucleotides which is a right arm end-sequence of YAC 942_c_3;

FIGURE 9 shows a sequence of nucleotides which is a left arm end-sequence of YAC 961_h_9;

FIGURE 10 shows a sequence of nucleotides which is a right arm end-sequence of YAC 961_h_9;

FIGURE 11 shows a sequence of nucleotides which is a left arm end-sequence of YAC 907_e_1;

FIGURE 12 shows a pedigree of family MAD31;

FIGURE 13 shows the haplotype analysis for family MAD13. Affected individuals are represented by filled diamonds, open diamonds represent individuals who were asymptomatic at the last psychiatric evaluation. Dark gray bars represent markers for which it cannot be deduced if they are recombinant; and

FIGURE 14 shows the YAC contig map of the region of human chromosome 18 between the polymorphic markers D18560 and D18561. Black lines represent positive hits. YACs are not drawn to scale.

FIGURE 15 shows (a) a CAG repeat (in bold) and

surrounding nucleotide sequence isolated from YAC 961_h_9. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 16 shows (a) a CAG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 766_f_12. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 17 shows (a) a CAG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 766_f_12. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

FIGURE 18 shows (a) a CTG repeat (in bold) and surrounding nucleotide sequence isolated from YAC 907_e_1. The sequence in italics is derived from End Rescue of the fragmented YAC. (b) PCR primers that can be used to determine the extent of trinucleotide repeats in the sequence.

Experimental 1

(a) Family Data

Clinical diagnoses in MAD31, a Belgian family with a BPII proband were described in detail in De bruyn et al 1996. In that study only the 15 family members who

were informative for linkage analysis were selected for additional genotyping. The different clinical diagnoses in the family were as follows:
1 BPI, 2 BPII, 2UP, 4 Major depressive disorder (MDD), 1 SAm and 1 SAd.
The pedigree of the MAD31 family is shown in Figure 12.

(b) Genotyping of Family Members

All short tandem repeat (STR) genetic markers are di- or tetranucleotide repeat polymorphisms. Information concerning the genetic markers used in this study was obtained from several sources on the internet: Genome DataBase (GDB, http://gdbwww.gdb.org/), GenBank (http://www.ncbi.nlm.nih.gov/), Cooperative Human Linkage Center (CHLC, http://www.chlc.org/), Eccles Institute of Human Genetics (EIHG, http://www.genetics.utah.edu/) and Généthon (http://www.genethon.fr/). Standard PCR was performed in a 25 μl volume containing 100 ng genomic DNA, 200 mM of each dNTP, 1.25 mM $MgCl_2$, 30 pmol of each primer and 0.2 units Goldstar DNA polymerase (Eurogentec). One primer was end-labelled before PCR with [gamma-^{32}P]ATP and T4 polynucleotide kinase. After an initial denaturation step at 94°C for 2 min, 27 cycles were performed at 94°C for 1 min, at the appropriate annealing temperature for 1.5 min and extension at 72°C for 2 min. Finally, an additional elongation step was performed at 72°C for 5 min. PCR products were detected by electrophoresis on a 6% denaturing polyacrylamide gel and by exposure to an X-ray sensitive film. Successfully analysed STSs, STRs and ESTs covering the refined candidate region are fully described herein on pages 36 to 54.

(c) Lod score analysis.

Two-point lod scores were calculated for 3 different disease models using Fastlink 2.2. (Cottingham et al. 1993). For all models, a disease gene frequency of 1% and a phenocopy rate of 1/1000 was used. Model 1 included all patients and unaffected individuals with the latter individuals being assigned to a disease penetrance class depending on their age at examination. The 9 age-dependent penetrance classes as described by De bruyn et al (1996) were multiplied by a factor 0.7 corresponding to a reduction of the maximal penetrance of 99% to 70% for individuals older than 60 years (Ott 1991). Model 2 is similar to model 1, but patients were assigned a diagnostic stability score, calculated based on clinical data such as the number of episodes, the number of symptoms during the worst episode and history of treatment (Rice et al. 1987, De bruyn et al. 1996). Model 3 is as model 1 but includes only patients.

(d) Construction of the YAC contig - protocols

Growing of YACs and extraction of YAC DNA was done according to standard protocols (Silverman, 1995). For the construction of the YAC-contig spanning the chromosome 18q candidate region, the data of the physical map based on sequence tagged sites (STSs) (Hudson et al. 1995) was consulted on the Whitehead Institute (WI) Internet site (http://www-genome.wi.mit.edu/). CEPH mega-YACs were obtained from the YAC Screening Centre Leiden (YSCL, the Netherlands) and from CEPH (Paris, France). The YACs were analyzed for the presence of STSs and STRs, previously located between D18S51 and D18S61, by

touchdown PCR amplification. Information on the STSs/STRs was obtained from the WI, GDB, Généthon, CHLC and GenBank sites on the Internet. Thirty PCR cycles consisted of: denaturation at 94°C for 1 min, annealing (2 cycles for each temperature) starting from 65°C and decreasing to 51°C for 1.5 min and extension at 72°C for 2 min. This was followed by 10 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1.5 min and extension at 72°C for 2 min. A final extension step was performed for 10 min at 72°C. Amplified products were visualised by electrophoresis on a 1% TBE agarose gel and ethidium bromide staining.

(e) Ordering of the STR markers.

Twelve STR markers, previously located between D18S51 and D18S61, were tested for cosegregation with bipolar disease in family MAD31. The parental haplotypes were reconstructed from genotype information of the siblings in family MAD31 and minimalizing the number of possible recombinants. The result of this analysis is shown in Figure 13. The father was not informative for 3 markers, the mother was not informative for 5 markers. Haplotypes in family MAD31 suggested the following order for the STR markers analysed: cen-[S51-S68-S346]-[S55-S969-S1113-S483-S465]-[S876-S477]-S979-[S466-S817-S61]-tel. The order relative to each other of the markers between brackets could not be inferred from our haplotype data. The marker order in family MAD31 was compared with the marker order obtained using different mapping techniques and the results shown in Table 1 below.

Table 1. Comparison of the order of the markers within the 18q candidate region for bipolar disorder, among several maps.

Marker*	Genetic maps		Radiation hybrid map
	Généthon	Marshfield	(Giacalone et al. 1996)
D18S51		(-)3.4cM	(-)27.9 cR
D18S68	0 cM	0 cM	0 cR
D18S346		5.3 cM	52.2 cR
D18S55	0.1 cM	0 cM	72.5 cR
D18S969		0.6 cM	
D18S1113	0.7 cM		
D18S483	2.5 cM	3.2 cM	88 cR
D18S465	4.5 cM	5.3 cM	101.3 cR
D18S876			
D18S477	4.4 cM	5.3 cM	166.4 cR
D18S979		8.9 cM	
D18S466	7.6 cM	11.1 cM	212.4 cR
D18S61	8.4 cM	11.8 cM	249.5 cR
D18S817		5.3 cM	260.6 cR

* Order according to haplotyping results in family MAD31.

(-) Marker is located proximal of D18S68.

D18S68, common to all 3 maps, was taken as the map anchor point, and the genetic distance in cM or cR of the other markers relative to D18S68 are given. The marker order is in good agreement with the order of the markers on the recently published chromosome 18 radiation hybrid map (Giacalone et al. (1996) Genomics 37:9-18) and the WI YAC-contig map (http://www-genome.wi.mit.edu/). However, a few discrepancies with other maps were observed. The only discrepancy with the Généthon genetic map is the reversed order of D18S465 and D18S477. Two discrepancies were observed with the Marshfield map (http://www.marshmed.org/genetics/). The present inventors mapped D18S346 above D18S55 based on maternal haplotypes, but on the Marshfield maps D18S346 is located between D18S483 and D18S979. The inventors also placed D18S817 below D18S979, but on the Marshfield map this marker is located between D18S465 and D18S477. However, the location of D18S346 and D18S817 is in agreement with the chromosome 18 radiation hybrid map of Giacalone et al. (1996). One discrepancy was also observed with the WI radiation hybrid map (http://www-genome.wi.mit.edu/), in which D18S68 was located below D18S465. However, the inventors as well as other maps placed this marker above D18S55.

(f) <u>Lod score analysis and refinement of the candidate region.</u>

Lod score analysis gave positive results with all markers, confirming the previous observation that 18q21.33-q23 is implicated in BP disease, at least in family MAD31 (De bruyn et al. 1996). Summary statistics of the lod score analysis under all models

are given in table 2 below.

Table 2. Summary statistics of the two-point lod scores in MAD31.

Marker	Model 1			Model 2			Model 3		
	Z at θ=0.0	Zmax	θmax	Z at θ=0.0	Zmax	θmax	Z at θ=0.0	Zmax	θmax
D18S51	-0.19	0.73	0.1	0.94	0.94	0.01	0.08	0.54	0.1
D18S68	-0.19	0.73	0.1	0.94	0.94	0.01	0.07	0.55	0.1
D18S346	-0.19	0.73	0.1	0.94	0.94	0.01	0.07	0.55	0.1
D18969	1.40	1.40	0.0	1.27	1.27	0.0	1.20	1.20	0.0
D18S1113	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S876	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S477	2.01	2.01	0.0	1.87	1.87	0.0	1.77	1.77	0.0
D18S979	-0.18	0.77	0.1	1.08	1.08	0.0	0.08	0.54	0.0
D18S817	-0.19	0.73	0.1	1.08	1.08	0.0	0.07	0.55	0.1
D18S61	-0.21	0.73	0.1	1.08	1.08	0.0	0.07	0.54	0.1

D18S55, D18S483, D18S465 and D18S466 were not informative.

The highest two-point lod score (+2.01 at θ=0.0) was obtained with markers D18S1113, D18S876 and D18S477 under model 1 in the absence of recombinants (table 2). In model 1, all individuals with a BP spectrum disorder are considered affected and fully contributing to the linkage analysis.
Before the fine mapping the candidate region was flanked by D18S51 and D18S61, which are separated by a genetic distance of 15.2 cM on the Marshfield map or 13.1 cM on the Généthon map. The informative recombinants with D18S51 and D18S61 were observed in 2 affected individuals (II.10 and II.11 in Fig. 13). However, since no other markers were tested within the candidate region it was not known whether these individuals actually shared a region identical-by-descent (IBD). The additional genetic mapping data now indicate that all affected individuals are sharing alleles at D18S969, D18S1113, D18S876 and D18S477 (Fig. 13, boxed haplotype). Also, alleles from markers D18S483 and D18S465 are probably IBD, but these markers were not informative in the affected parent I.1. Obligate recombinants were observed with the STR markers D18S68, D18S346, D18S979 and D18S817 (Table 2, fig. 13) Since discrepancies between different maps were observed for the locations of D18S346 and D18S817, the present inventors used D18S68 and D18S979 to redefine the candidate region for BP disease. The genetic distance between these 2 markers is 8.9 cM based on the Marshfield genetic map (http.//www.marshmed.org/genetics/).

(g) Construction of the YAC contig.

According to the WI integrated map 56 CEPH megaYACs are located in the initial candidate region

contained between D18S51 and D18S61 (Chumakov et al. (1995) Nature **377** Suppl., De bruyn et al. (1996)). From these YACs, those were selected that were located in the region between D18S60 and D18S61. D18S51 is not presented on the WI map, but is located close to D18S60 according to the Marshfield genetic map (http.//www.marshmed.org/genetics/). To limit the number of potential chimaeric YACs,YACs were eliminated that were also positive for non-chromosome 18 STSs. As such, 25 YACs were selected (see Figure 14), and placed in a contig based on the technique of YAC contig mapping, i.e. sequences from sequence tagged sites (STSs), simple tandem repeats (STRs) and expressed sequenced tags (ESTs), known to map between D18S60 and D18S61, were amplified by PCR on the DNA from the YAC clones. The STS, STR and EST sequences used, are described from page 36 to 54. Positive YAC clones were assembled in a YAC contig map (Figure 14).

Three gaps remained in the YAC contig, of which one, between D18S876 and GCT3G01, was located in the refined candidate region. To close the gap between D18S876 and GCT3G01, 14 YAC clones (Table 3, on page 62) were further analysed. End fragments from YAC clones 766_f_12 (SV11R), 752_g_8 (SV31L), 942_c_3 (SV10R) were obtained and sequenced (see pages 55-61). Primers from these three sequences were selected, and DNA of each of the 14 YAC clones was amplified by PCR. As indicated in Table 3, overlaps were obtained between 7 YAC clones on the centromeric side, and two YAC clones on the telomeric site (717_d_3 and 907_e_1).

The final YAC contig is shown in Figure 14. In the figure, only the YAC clones which rendered unambiguous hits with the chromosome 18 STSs, STRs and ESTs are shown. In a few cases, weak positive signals were also obtained with some of the YAC clones, which

likely represent false positive results. However, these signals did not influence the alignment of the YAC clones in the contig. Although, all YACs known to map in the region were tested as well as all available STSs/STRs, initially, the gap in the YAC contig was not closed. However, this was subsequently achieved by determining the end-sequences of the eight selected YACs (see below). The order of the markers provided by the YAC contig map is in complete agreement with the marker order provided by the WI map which integrates information from the genetic map, the radiation hybrid map and the STS YAC contig map (Hudson et al. 1995). Also, the YAC contig map confirms the order of the STR markers as suggested by the haplotype analysis in family MAD31. Moreover, the YAC contig map provides additional information on the relative order of the STR markers. For example, D18S55 is present in YAC 931_g_10 but not in 931_f_1 (Fig.14), separating D18S55 from its cluster [S55-S969-S1113-S483-S465] obtained by haplotype analysis in family MAD31. The centromeric location of D18S55 is defined by the STS/STR content of surrounding YACs (Fig. 14). If we combine the haplotype data and the YAC contig map the following order of STR markers is obtained: cen-[S51-S68-S346]-S55-[S969-S1113]-[S483-S465]-S876-S477-S979-S466-[S817-S61]-tel.

Out of the 25 YAC clones spanning the whole contig, seven YAC clones were selected in order to identify the minimal tiling path (Table 4). These 7 YAC clones cover the whole refined chromosome 18 region. Furthermore, YAC clones should preferably be non-chimeric, i.e. they should only contain fragments from human chromosome 18. In order to examine for the presence of chimerism, both ends of these YACs were subcloned and sequenced (pages 55 to 61). For each of

the sequences, primers were obtained, and DNA from a monochromosomal mapping panel was amplified by PCR using these primers. As indicated on pages 55 to 61, some of the YAC clones contained fragments from other chromosomes, apart from human chromosome 18.

Three YAC clones were then selected comprising the minimum tiling path (Table 5). These three YAC clones were stable as determined by pulsed field gel electrophoresis and their seizes correspond well to the published sizes. These YAC clones were transferred to other host yeast strains for restriction mapping, and are the subject to further subcloning.

Description of the succesfully analysed STSs, STRs and ESTs covering the refined candidate region.

Explanations:

- STS: Sequence Tagged Site
- STR: Simple Tandem Repeat
- EST: Expressed Sequence Tag

These markers are ordered from the centromere to the telomere. Only the markers that were effectively tested and that worked on the YACs are given.

List:

1. D18S60:

Database ID: AFM178XE3 (Also known as 178xe3, Z16781, D18S60)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = CCTGGCTCACCTGGCA
Right = TTGTAGCATCGTTGTAATGTTCC
Product Length = 157
Review complete sequence:
AGCTATCCTGGCTCACCTGGCAAAAATACAGTGTATACACACACACACAC
ACACACACACACACACACAGAGTGTNTTANTNATTCCAGCAAATAATATTA
CATATAAAAGATCTAATTGGTTCATCATGTAAATTTAGTAGGAACATTACA
ACGATGCTACAAGANTTTATCCAAAACTGAGATTTCCTTAGAATATCTGTT
AAAAGTAATTTTATTCAGTTAATAGAAATTCTATTGAAAACATCAAACTTAT
AAAGCT

Genbank ID: Z16781
Description: H. sapiens (D18S60) DNA segment containing (CA) repeat; clone
Search for GDB entry

2. WI-9222:

Database ID: UTR-03540 (Also known as G06101, D18S1033, 9222, X63657)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:

Left = GATCCCATAAAGCTACGAGGG
Right = GAGTCTAAAGACAAGAAAGCATTGC
Product Length = 99
Review complete sequence:

TCTTCTTACCCCTTGGAAGAAGACTGTTTCCAAATAATTTGAACAGCTTG
CTGCTAAATGGGACCCAATTTTTGGCCTATAGACACTTATGTATTGTTTTC
GAATACGTCAGATTGGACCAGTGCTCTTCAGGAATGTGGCTGCAAGCAA
GGGGCTAGAAGTTCACCTCCTGACAGTATTATTAATACTATGCAAATATG
GAATAGGAGACCATTTGATTTTCTAGGCTTTGTGGTAGAGAGGTGAAGG
TATGAGAATTAATAGCGTGTGAACAAAGTAAAGAACAGGATTCCAGAATG
ATCATTAAATTTGTTTCTATTTATTCTTTTTTGCCCCCCTAGAGATTAAGTC
CAGAAATGTACTTTCTGGCACATAAAGAAATCTTGAGGACTTTGTTTAAAC
CTTCCATAAAAAAACAATTTTCGGTTTCTCGGGTNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT
TCTTTCTTTGTGTATTTTATTCAAGATGAGTTGGACCCATTGCCAGTGAGT
CTGAATGTCACTGACAGCCCTGTGTTGTGCTCAGGACTCACTCTGCTGC
TGGTGGAAACTCATGGCTTCTCTCTCTCTTTGATCCCATAAAGCTACGAG
GGGGACGGGAGAGGGCAGTGCAATGGGAAGTAAAGAGATATTTTCCAG
TAGGAAAAGCAATGCTTTCTTGTCTTTAGACTCAAATGCTTAGGGAACGT
TTCATTTCTCATTCATGGGGAAAGGCAGCCTCCTTAAATGTTTTCTGAAG
AGCGGTAAAATCTAGAAGCTTAAGAATTTACAGTTCCTTCAATAACCATGA
TGACCTGAAGTTCACCTATCCCATTTTAGCATCTACTTGTTTTTCCCATCT
CTTCCTTTCCAATTTTGCTTATACTGCTGTAATATTTTTGTNNNNNNNNNN
NNNNNNNNNNNNNNGACCAGCTAAAATTTTCGACTTGACTTTTTAACTT
AACTCATGAATTAATTAAAGCAAATGAAAAAATTAAAAAGTGTGACTTTTT
CTCGGAGCATATATGTAGCTTTTAGGAAAGGCTGATGATGGTATAAAGTT
TGCTCATTAAGAAAAAAAGACAAGGCTGATTTTGAAGAGAGTTGCTTTTG
AAATAAAATGATCA

Genbank ID: X63657
Description: H.sapiens fvt1 mRNA
Search for GDB entry

3. WI-7336:

Database ID: UTR-04664 (Also known as PI5, G00-679-135, G06527, 7336, U04313)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = AGACATTCTCGCTTCCCTGA
Right = AATTTTGACCCCTTATGGGC
Product Length = 332
Review complete sequence:

TAAGTGGCATAGCCCATGTTAAGTCCTCCCTGACTTTTCTGTGGATGCCG
ATTTCTGTAAACTCTGCATCCAGAGATTCATTTTCTAGATACAATAAATTG
CTAATGTTGCTGGATCAGGAAGCCGCCAGTACTTGTCATATGTAGCCTTC

ACACAGATAGACCNNNNNNNNNNCCAATTCTATCTTTTGTTTCCTTTTTT
CCCATAAGACAATGACATACGCTTTTAATGAAAAGGAATCACGTTAGAGG
AAAAATATTTATTCATTATTTGTCAAATTGTCCGGGGTAGTTGGCAGAAAT
ACAGTCTTCCACAAAGAAAATTCCTATAAGGAAGATTTGGAAGCTCTTCT
TCCCAGCACTATGCTTTCCTTCTTTGGGATAGAGAATGTTCCAGACATTC
TCGCTTCCCTGAAAGACTGAAGAAAGTGTAGTGCATGGGACCCACGAAA
CTGCCCTGGCTCCAGTGAAACTTGGGCACATGCTCAGGCTACTATAGGT
CCAGAAGTCCTTATGTTAAGCCCTGGCAGGCAGGTGTTTATTAAAATTCT
GAATTTTGGGGATTTTCAAAAGATAATATTTTACATACACTGTATGTTATA
GAACTTCATGGATCAGATCTGGGGCAGCAACCTATAAATCAACACCTTAA
TATGCTGCAACAAAATGTAGAATATTCAGACAAAATGGATACATAAAGACT
AAGTAGCCCATAAGGGGTCAAAATTTGCTGCCAAATGCGTATGCCACCA
ACTTACAAAAACACTTCGTTCGCAGAGCTTTTCAGATTGTGGAATGTTGG
ATAAGGAATTATAGACCTCTAGTAGCTGAAATGCAAGACCCCAAGAGGAA
GTTCAGATCTTAATATAAATTCACTTTCATTTTTGATAGCTGTCCCATCTG
GTCATGTGGTTGGCACTAGACTGGTGGCAGGGGCTTCTAGCTGACTCG
CACAGGGATTCTCACAATAGCCGATATCAGAATTTGTGTTGAAGGAACTT
GTCTCTTCATCTAATATGATAGCGGGAAAAGGAGAGGAAACTACTGCCTT
TAGAAAATATAAGTAAAGTGATTAAAGTGCTCACGTTACCTTGACACATAG
TTTTTCAGTCTATGGGTTTAGTTACTTTAGATGGCAAGCATGTAACTTATA
TTAATAGTAATTTGTAAAGTTGGGTGGATAAGCTATCCCTGTTGCCGGTT
CATGGATTACTTCTCTATAAAAAATATATATTTACCAAAAAATTTTGTGACA
TTCCTTCTCCCATCTCTTCCTTGACATGCATTGTAAATAGGTTCTTCTTGT
TCTGAGATTCAATATTGAATTTCTCCTATGCTATTGACAATAAAATATTATT
GAACTACC

Genbank ID: G06527
Description: WICGR: Random genome wide STSs

4. WI-8145:

Database ID: EST102441 (Also known as D18S1234, G00-677-827, G06845, 8145, T49159)
Source: WICGR: STSs derived from dbEST sequences
Chromosome: Chr18

Primers:
Left = GAAATGCACATAACATATATTTGCC
Right = TGCTCACTGCCTATTTAATGTAGC
Product Length = 184
Review complete sequence:

GTTGTTTGGANGCAGGTTTATTTATTATATACTTGCAATTGAATATAAGAT
ACAGACATATATATGTGTTATGTATTTCTAGAAATGCACATAACATATATTT
GCCTATTGTTTAATGTTTTTTCCAGANATTTATTACAGAAGGGCATGGAG
GGATACCTACTTATTCTTCATTATGAGAACAATTAAAGGCATTTATTAGAT
AGGAAATTAACAGANCATCTGCTTCTATAACTTTATTAGCTACATTAAATA
GGCAGTGAGCANTAATTTAAAANCTCACCATTATATAAANTANTAAATACC
AAAGTAAAAG

_____: left and right primer

PCR Conditions

Genbank ID: T49159
Description: yb09e07.s1 Homo sapiens cDNA clone 70692 3' similar to gb:J02685
UniGene Cluster Description: Human mRNA for Arg-Serpin (plasminogen activator-inhibitor 2, PAI-2) Search for GDB entry

5. WI-7061:

Database ID: UTR-02902 (Also known as PAI2, G00-678-979, G06377, 7061, M18082)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = TGCTCTTCTGAACAACTTCTGC
Right = ATAGAAGGGCATGGAGGGAT
Product Length = 338
Review complete sequence:

AACTAAGCGTGCTGCTTCTGCAAAAGATTTTTGTAGATGAGCTGTGTGCC
TCAGAATTGCTATTTCAAATTGCCAAAAATTTAGAGATGTTTTCTACATAT
TTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAATAAAAACACAGAAAT
AATTAGACAATTGTCTATTATAACATGACAACCCTATTAATCATTTGGTCT
TCTAAAATGGGATCATGCCCATTTAGATTTTCCTTACTATCAGTTTATTTT
TATAACATTAACTTTTACTTTGTTATTTATTATTTTATATAATGGTGAGTTTT
AAATTATTGCTCACTGCCTATTTAATGTAGCTAATAAAGTTATAGAAGCAG
ATGATCTGTTAATTTCCTATCTAATAAATGCCTTTAATTGTTCTCATAATGA
AGAATAAGTAGGTATCCCTCCATGCCCTTCTATAATAAATATCTGGAAAAA
ACATTAAACAATAGGCAAATATATGTTATGTGCATTTCTAGAAATACATAA
CACATATATATGTCTGTATCTTATATTCAATTGCAAGTATATAATAAATAAA
CCTGCTTCCAAACAACNNNNNNNNNNNNNNGGAATTC

PCR Conditions

Genbank ID: G06377
Description: WICGR: Random genome wide STSs

6. D18S68:

Database ID: AFM248YB9 (Also known as 248yb9, Z17122, D18S68)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATGGGAGACGTAATACACCC
Right = ATGCTGCTGGTCTGAGG
Product Length = 285
Review complete sequence:

AAAGAGTTGGGGTTGTGAATTCCCACACCAGTCAACTATTGGCTATGGG
CTTACCATGGGAGACGTAATACACCCGGNACTTCCAACTCACATACCAG
AGACATGGCTCTAGCACCCAATGGAAATATGCTGAATGTTGCAGGTGCA
AGACAGCAACAAAGCAGACAGAGGCACATAGACAAGGCACCAACAGTGT
CCACTATACCCTGACAGTGTGGAAAGTTGTAGATAGGATGAAGAGAAAG
AATACACACACACACACACACACACACACACACACACACACACACACACA
CGGTAGANACTTACTACNCAAAGTGTGANCCTCAGACCAGCAGCATCTG
GCNAAATGGTGATCTATCACCTTCCAG

Genbank ID: Z17122
Description: H. sapiens (D18S68) DNA segment containing (CA) repeat; clone

7. WI-3170:

Database ID: MR3726 (Also known as D18S1037, G04207, HALd22f2, 3170)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TGTGCTACTGATTAAGGTAAAGGC
Right = TGCTTCTTCAATTTGTAGAGTTGG
Product Length = 156
Review complete sequence

CTGAGACAAGGCAGGCAAACAACCTCTAAAAATCTACAATTGGTGATTGG
TGTGCTACTGATTAAGGTAAAGGCACAGAATTATACATCCAGGTTNCTAT
TACTTATGGCAGACTCAGGACCCAGGTTNAGAGACCACTGGCCTTAAGA
AAAAAAATGGGGTTCCTGATTTCTGGATAATAATCCAACTCTACAAATTGA
AGAAGCAACATACCCTCTTTGTTA

Genbank ID: G04207
Description: WICGR: Random genome wide STSs

8. WI-5654:

Database ID: MR10908 (Also known as D18S1259, G00-678-695, G05278, 5654)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = CTTAATGAAAACAATGCCAGAGC
Right = TGCAAAATGTGGAATAATCTGG
Product Length = 149
Review complete sequence:

CTACAAAATGCATGTGGCTTTGGCTTTGAAATAGTACACCCTATCAAAGA
CTAAATTTTCTTAATGAAAACAATGCCAGAGCTTTTTTCATGATATTTTGTT
TTTAGAGATGGGGAACAATCTGGACGTTGTTTCCTTATCTGGGTGGTAAT
CGAGGCTTAGCAATTTCCCACAGCGTTACACAAATCCAGATTATTCCACA
TTTTGCAAATA

Genbank ID: G05278
Description: WICGR: Random genome wide STSs

9. D18S55:

Database ID: AFM122XC1 (Also known as 122xc1, Z16621, D18S55, GC378-D18S55)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18
Primers:
Left = GGGAAGTCAAATGCAAAATC
Right = AGCTTCTGAGTAATCTTATGCTGTG
Product Length = 143
Review complete sequence:
AGCTGAACATGCCTTTTCATGGAGCAGTTTCNAAATACACTTTTGGTACA
ATCTGCAGGTGGATATTTGGAGCTCAGGAGTTTGAGACCAGCCTGGGCA
ACATGGTGAAATCCCGTCTCTACTAAAATACAAAAAATTAGCCAGGTGTG
GCGGCATGTGCCTGTAGNCCCAGGATGGATTGAGTGGGTGAGATATGG
AATAAGTGGTGGGAAGTCAAATGCAAAATCAATTCAGTTTGTCAATATTG
ATTCTCTATTCTAGCCTGGCGTGGTTTTTCCTCGTCACACACACACACAC
ACACACACACACACACACACACACACACACACACACACAGCATAAGATTACTCAGA
AGCT

Genbank ID: Z16621
Description: H. sapiens (D18S55) DNA segment containing (CA) repeat; clone

10. D18S969:

Database ID: GATA-P18099 (Also known as G08003, CHLC.GATA69F01, CHLC.GATA69F01.P18099)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AACAAGTGTGTATGGGGGTG

Right = CATATTCACCCAGTTTGTTGC
Product Length = 365
Review complete sequence:

CAGGGAAATGCAAATCAAAACCACAATGAGTTATCTCCTCATACCTTTAAT
GATGGCTAATATTAAACAAGAGATAACAAGTGTGTATGGGGGTGTGGAG
AAAAGAGAATGTNCGAACACTCTTGGTTGAAATATAAGTTGGTAGANCCA
TTATGCAAAACAGTATGAATCTTTATCAGTATAANATTAGGACCTNGCATA
TGATCNCAGCAATCNCCACNTCTGNGNGATCNCACNCNCTATCTCTCTAT
ATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT
ATCTGTCTGTCTATCATCTATCTATCTTCTATCTATCTATCTATCTTTCTAT
CTATCTATCTGTCTATCTATNCCGGAATATTTTTCAGCCATNNAAATAAGG
AAGTCCTGCTATTTGCAACAAACTGGGTGAATATGGAGAACGTTATGCTA
AATGCAATATGCTAAAGACAGACACAGAAAGACAAGTATGACCTCACTTA
TATGTGGAAACTGAAAAAGCCATACTCATTACAGCAAAGAGTAGAATGTT
GGTTACCAGGGGCAAAGAGGGTAGAAATGAGGGGAGTGAGAAAATGTC
AATCAAAGTGTAAGAATGTTATAACATAAATAAATTCATAGAG

Genbank ID: G08003
Description: human STS CHLC.GATA69F01.P18099 clone GATA69F01.

11. D18S1113:

Database ID: AFM200VG9 (Also known as D18S1113, 200vg9, w2403)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = GTTGACTCAAGTCCAAACCTG
Right = CAAAGACATTGTAGACGTTCTCTG
Product Length = 207
Review complete sequence:
AGCTGCATATAAAACTATTCCATTTCACATTTTTGAAGACATTTGTAGCCA
TGATACTTTGCTGTTGTCTGTGGGCCACCTCTTTTTGAAGTGTGTAGTTA
ACTGTGCTCCTGTAATCTGTTGTCTGTTGACTCAAGTCCAAACCTGTTCT
GCGTGGCATGTTTCTNCAACTTGATGTGATGCTATTTATCACTTTCTTTGA
AGTTAAGTCTCTATGTCTTTGTATTCTTTCTGTGTACCCAGGGATATGTTT
GTGCATGCACACGCATAAACACACACACACACACACACACACACAGAGA
CAGAGACAGAGAACGTCTACAATGTCTTTGTGAG

12. D18S868:

Database ID: GATA-D18S868 (Also known as G09150, CHLC.GATA3E12, CHLC.GATA3E12.496, CHLC.496, D18S868)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AGCCAATACCTTGTAGTAAATATCC
Right = GATTCTCCAGACAAATAATCCC
Product Length = 189
Review complete sequence:
GAGTGAGCCAATACCTTGTAGTAAATATCCATCTATCTTTGATGTATCTAT
GTATCTATCTTTGTATCTATATGTCTATGTATCTATGTATGTATGTATCTAT
CTATCATCTATCTATCTATCATCTATCTATCTATCTATCTATCTATCTATCT
ATCTATCTATATCCNTTTGGGATTATTTGTCTGGAGAATCCTGATTAACAT
AGTCTGCTAACTTTTATCTGTATCTCCTATGTGTATGCTTCTCCTTCTTCC
TGTCTCTCTCTCTTCTTTGTCCTCATTTAANCTCCTTTCCTGGGNATATTG
GNAATTTTGATTGGANTCTGGACANTGTAGGAGTAAAAATTT

Genbank ID: G09150
Description: human STS CHLC.GATA3E12.P6553 clone GATA3E12.

13. WI-9959:

Database ID: MR12816 (Also known as D18S1251, G00-678-524, G05488, 9959)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TGCCAACAGCAGTCAAGC
Right = AGCACCTGCAGCAGTAATAGC
Product Length = 110
Review complete sequence:
ctgttttatttgaaaaaaaaaatctgtctccaagaagaaaagttcattctACCTGTTGCCAACAGC
AGTCAAGCGGACATGTTTAAAATTTTTTAAAAAAGTATTTTTTTTTCCAACT
GGNGTTTAATAGCCTCATTTTGGCTTTTGCTATTACTGCTGCAGGTGCTT
TNATTTTTTTCCTCTGCATTATAATTAC

Genbank ID: G05488
Description: WICGR: Random genome wide STSs
Search for GDB entry

14. D18S537:

Database ID: CHLC.GATA2E06.13 (Also known as CHLC.13, GATA2E06, D18S537, GATA-D18S537)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCCATCTATCTTTGATGTATCTATG
Right = AGTTAGCAGACTATGTTAATCAGGA
Product Length = 191

Review complete sequence:
AAAGCTGAGTGAGCCAATACCTTGTAGTAAATATCCATCTATCTTTGATGT
ATCTATGTATCTATCTTTGTATCTATATGTCTATGTATCTATGTATGTATGT
ATCTATCTATCATCTATCTATCTATCATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATATCCNTTNGGTATTATTNGTCTGGNGAATCCTGAT
TAACATAGTCTGCTAACTTNTATCTGTATCTNCTATGTGTATGCTTCTNCT
TCTTCCTGTCTCTCTCTCTGCTTTGTCCTCAATTNAAATCTCC

Genbank ID: G07990
Description: human STS CHLC.GATA2E06.P6006 clone GATA2E06.
Search for GDB entry

15. D18S483:

Database ID: AFM324WC9 (Also known as 324wc9, Z24399, D18S483)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = TTCTGCACAATTTCAATAGATTC
Right = GAACTGAGCAAACGAGTATGA
Product Length = 214
Review complete sequence:
AGCTCTGCTGGAAGAGCAGGGCTGTTTTCTGCACAATTTCAATAGATTCC
CCTACCCTGGGTTTTTCAGTAGATAGATAGATAGATGATAGATAGGTAGA
TAGATAGATAGATAGATAGATAGATAGATAGATAGATGATAGATAGATTTT
ATATATAGTATATAAAATCTACACACACACACACACACACACACACACATA
TTTGCCTTTCCTTGACTATCATACTCGTTTGCTCAGTTCTTTTTTTTTTAA
ATTTTTGTTTGTAAATCCAAAATGCTT

Genbank ID: Z24399
Description: H. sapiens (D18S483) DNA segment containing (CA) repeat; clone
Search for GDB entry

16. D18S465:

Database ID: AFM250YH1 (Also known as 260yh1, Z23850, D18S465)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATATTCCCCTATGGAAGTACAG
Right = AAAGTTAATTTTCAGGCACTCT
Product Length = 232
Review complete sequence:
AGCTCTGTCCCTCTAGAGAACGCTGACTAATATATTCCCCTATGGAAGTA
CAGATGGTTTTTTNTAAAATAAATTTATCTGATTGTGATGAGATAATCATCA

TTTTTATGTTCAGTGTTTTTCTAAATTTTTTATTGTTATTGTTTTTATACTCT
AAATGGTTTTTAAATATGCACATATGTGCATATTTTACACACACACACACA
CACACACACACTCTCTTTATTTAGAAGCATTATAGATAGAGTGCCTGAAAA
TTAACTTTTAACCNAAGAAAAGACAATAAGGAACAATAGGGAAGTTATCC
TTTGCTAAGGGTATGGAAAATATTCACATATTATTTATAACANGTTAAACC
AAGTCATGCTTGANTATAATAGCT

Genbank ID: Z23850
Description: H. sapiens (D18S465) DNA segment containing (CA) repeat; clone
Search for GDB entry

17. D18S968:

Database ID: GATA-P34272 (Also known as G10262, CHLC.GATA117C05, CHLC.GATA117C05.P34272)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = GAAATTAACCAGACACTCCTAACC
Right = CTTAGAATTGCCTTTGCTGC
Product Length = 147
Review complete sequence:
GAATAAAAATATGAGGTATTAGAAATTTACAGATAGGAAGAAATTAACCAG
ACACTCCTAACCACCGATNAGTTTAAAGAGGAGATAGATAGATAGATGAT
AGATAGATAGATAGATAGATAGATACCACTGAAAATGCAANCACAAATTA
GCAGATTATATGTGATGCAGCAAAGGCAATTCTAAGTAGATTCTAACTGC
TACATTGATAGCAGTACCCACTGACATTACCGGAAAGGATGGTATCCATA
ACCACCTACCTATATACCTCCGCAGCTGGANATTAGGNTTAAGCTTCTTN
GGGCNCCTGGCGGCCCCNNTTGTGGTCCCCGGTNGGNCCCCGNTTNN
GNNTNGCTNNGNTTNCNTTGGNGNCCCCCNNTNGGTTTNNGGNNNNNT
NNNNNTNGNNNNNTTNCCCNNNNNNNNTNTNNNNCNNNNNNNNTNNN
NNNNNNNNNNGGNNNNNGGGN

Genbank ID: G10262
Description: human STS CHLC.GATA117C05.P34272 clone GATA117C05.

18. GATA-P6051:

Database ID: GATA-P6051 (Also known as CHLC.GATA3E08, CHLC.GATA3E08.P6051)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = GCAACAACCCTAATGAGTATACG

Right = GAGTCTCACCAGGGCTTACA
Product Length = 149
Review complete sequence:
AAAGCTGTCTCCTTTTGTAAAGTGTGCTCAGAGGAATCTTTTTCAGTAAAT
AAAGTCTGCACCCAGACATCTCACTTTGTATACCACGGAGAATTTACCAT
GACTCTTCTCAGTGATAAACGTCAATATAGAATAATCAGGAGAAAAAGAG
AAATCCAGTAAAGAAATAAGTCTGTAGAAAGCAACAACCCTAATGAGTAT
ACGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATG
NATCTATCTATCTAACATTATATAAAATATATATTTCCTCCTGTATTGGGG
CCCTGTGTGTAAGCCCTGGTGAGACTCAAAAATTTGANTATTCCTNTTTN
T

Genbank ID: G09104
Description: human STS CHLC.GATA3E08.P6051 clone GATA3E08.

19. D18S875:

Database ID: GATA-D18S875 (Also known as G08001, CHLC.GATA52H04, D18S875)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCCTCTCATCTCGGATATGG
Right = AAGGCTTTCAGACTTACACTGG
Product Length = 394
Review complete sequence:
TTATTTATTCACTCATTCAATAAATATTTATGAATTTCCTTTAATGGCNANG
AAAGTATGTTTGGTACTGAATATGGTGAGCAAGATTTTCCTCTCATCTCG
GATATGGAAAGATCTTGGAAATCATTATACNTCATACTTACAATANGAAAG
AAGCTGAGCAATTTGAAAATCAACAATTTCTTTTGTACNTGTCAGAAAAGT
GAAGATATATTAATCAGGGTTCTTCAGAGAAACATAACCAATAGGNCACA
GNTCTATATGNCCNCNTTTATCTATCTATCTATCTATCTATCNCTATCTAT
CNCANACCNGGNGAANTNATNTTTGNGAGATTNATGCAAGNCTGAGAAA
NACCNAAGAANCTGCTCCCTGTNAAACTNGAGATNCAAGAANCTGAANA
GTATAGNTCCAGTCCNAAGTCTANAGACCTTAGAATTAGGAAAACTGATA
CTATAAATACCAGTGTAAGTCTGAAAGCCTTAAANACCANATAGTGCCAT
TGAAAGGGCAGAAGACTGATGTCCCAGTTCAAGCAGGCAAAGTTAGAGA
AGCCTTATTTTCTGCAACATTGTTCTATTCAGACCCTTNANANGATTGACN
ATGTCCACCCA

Genbank ID: G08001
Description: human STS CHLC.GATA52H04.P16177 clone GATA52H04.
Search for GDB entry

20. WI-2620:

Database ID: MR1436 (Also known as G03602, D18S890, HHAa12h3, 2620)
Source: WICGR: Random genome wide STSs

Chromosome: Chr18

Primers:
Left = TCTCCAAGCTATTGATTGGATAA
Right = TTAAGAGCCAATTTATATAAAAGCAGC
Product Length = 177
Review complete sequence:
CCCCTTTTGCCAACGCCATGCTTCACGTAGGGAGCCTGACATGCAGAAA
ACTCTCCAAGCTATTGATTGGATAAAGAGCCAGAGCTGACTGAATTCCAT
TCTTCTTGAGCCTCTCATTCTGTGTTTCTCGAATTTTTACCAAAGCATCTT
GACACACAAATATCTGACTCAAGGAAAAGGAAAAACAACTGCTTTTTCTC
CAGCTGCTTTTATATAAATTGGCTCTTAAACTTTCTAAGTTTATTATGGAT
A

Genbank ID: G03602
Description: WICGR: Random genome wide STSs
Search for GDB entry

21. WI-4211:

Database ID: MR6638 (Also known as G03617, D18S980, 4211)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = ATGCTTCAGGATGACGTAATACA
Right = AAATTCTCGCTGATTGGAGG
Product Length = 113
Review complete sequence:
CTAGTACCATAATCCCTTTTGGAATAAACCATCCCACCTTTAGTCAGANC
AGATGCTTCAGGATGACGTAATACATAATAAGCCTACTCAGTTCTACTCT
GGCTTTGTATGTCTTCAAAGTGATATTTTTTTAAGTATTACTTGTCCCTCC
AATCAGCGAGAATTT

Genbank ID: G03617
Description: WICGR: Random genome wide STSs
Search for GDB entry

22. D18S876:

Database ID: GATA-D18S876 (Also known as G09963, CHLC.GATA61E10, D18S876)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = TCAAACTTATAACTGCAGAGAACG
Right = ATGGTAAACCCTCCCCATTA

Product Length = 171
Review complete sequence:
AAGACTGCAATTACATTTGCATCAAACTTATAACTGCAGAGAACGTTGCC
CACTATTTTATACCACACAACAGTATTCTTAGCCAGATTACATCTATCTAT
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCATCTATCTAGC
TAGCTATCTATCTATAGAACTAATGGGGAGGGTTTACCATGTTTGGGTGA
ACCCAAACATTTTATGGNCAAGGGNTTGGAAAATTACCCTTATCTACAAC
TNTTNAACTTGTTTTGGTAGGNGTGNTAATTCCNTGGGNTTGGAANAACT
TTTGNAATTTCCTCNTTGTTTNTNATTNNNNATTNNTNNNCATTATTNTGG
GGTNTTCNGGGTGGAGGGCTNANTTTGGCCNCCCGGGTCCNNGGNGC
NAGTNGGNNNGGNTNNTNGGGTTTNCTTGGGAANCNTNCCNCCTNCNG
GGGNTTCANGGGNTTTTTNTTTNNTTG

Genbank ID: G09963
Description: human STS CHLC.GATA61E10.P17745 clone GATA61E10.
Search for GDB entry

23. GCT3G01:

Database ID: GCT-P10825 (Also known as G09484, CHLC.GCT3G01, CHLC.GCT3G01.P10825)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = CTTTGCAATCTTAGTTAATTGGC
Right = GAACTATGATATGGAGTAACAGCG
Product Length = 128
Review complete sequence:
AGATGTTTAACTTTGCAATCTTAGTTAATTGGCAGAAATGAAATTTAGTTT
CCACAACTTTTATTCGATATTAAAACACCACCACCATCAGCAGCAGCAGC
AGCAGCAGCAGCATCGCTGTTACTCCATATCATAGTTCAGAGCATTTAAA
GNGGTCAAAATATACAACTAGGCTGACACCNGNATAAGGTTTAATTTTAA
ACCNGNGGTCTNCCCTCTAAGGNGGNTTTTTTTTTCTTGNCNTGGCTTCT
TTTTCCNTTTGCTTTTGTAAAATATCAAGGNATTTTTGGGTTNTTCNTGGN
ANTTNNCNNANTNNTNNTTNNNCNCNCCCCCCNTTTGNGGCGGGGGTC
CCCNNNTTGCCCCGGGGTTGNGTGCAGTAGGGGGGTCNCGGGTNNNG
NAAGTTTNGGGGCCCT

Genbank ID: G09484
Description: human STS CHLC.GCT3G01.P10825 clone GCT3G01.

24. WI-528:

Database ID: MH232 (Also known as G03589, 528, D18S828)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = TTCTGCCTTTCCTGACTGTC
Right = TGTTTCCCATGTCTTGATGA
Product Length = 211
Review complete sequence:
CTACTAAGCAAATTCTGCTCAGCCTTCTGCCTTTCCTGACTGTCTTGTTG
GCCCTTCCCACTTTAAGGATGCCTGTTTAAGTAGCCACCTCTAATTAGGA
ATCTTCCCTTGTTCTTTCTCAGGAGGCTTAGACACTGTCAGTTTCCTGAA
GACAGAAAATAAGCCTGCATTATCCTAGTAGTGGATTCAAAACTAATTGT
GTCCTGAGTCTTTCAATCATCAAGACATGGGAAACACTCAACAG

Genbank ID: G03589
Description: WICGR: Random genome wide STSs
Search for GDB entry

25. WI-1783:

Database ID: MR432 (Also known as G03587, _shu_31.Seq, 1783, D18S824)
Source: WICGR: Random genome wide STSs
Chromosome: Chr18

Primers:
Left = CCAGTAATTAGACATTGACAGGTTC
Right = TTTTACTAGACAGGCTTGATAAACAA
Product Length = 305
Review complete sequence:
CCAGTAATTAGACATTGACAGGTTCCATACTAGTAATGTAGGGAATAGGG
CTGCTGCTTTTTGGGTTTCCTTGAGTATACTTTGTGCTGCATAAATATGG
CAATGGATAGTAAATAATTTGTATGCAGACCTTTAGTGTCGATTAACCTGT
GAATAAGGGAACAACAATCAAGGACAAAAATCAAAAGACTAATTCTCTAT
ACATTTTGAGCTTTTGTAAAAAAGTAAGATTAGCTGAATATATCTGAAAAA
TTTCTAATCTCCTTTACAATTTTTTAAATTGTTTATCAAGCCTGTCTAGTAA
AAATAATTCAGTTTCGGAATGTGG

Genbank ID: G03587
Description: WICGR: Random genome wide STSs
Search for GDB entry

26. D18S477:

Database ID: AFM301XF5 (Also known as 301xf5, Z24212, D18S477)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = GGACATCCTTGATTTGCTCATAA
Right = GATTGACTGAAAACAGGCACAT

Product Length = 243
Review complete sequence:
GGACATCCTTGATTTGCTCATAATACACTCATTCCTTTCACCATTGAGTGT
GCACATATTTCTCTGATTGGAAAGAACTACAGAGGAGGTTTTACNTTTTA
CTTTCCAGTTTGCTATTAAAGAGAGAAAACTAACAGAGNGAAATCAAGCA
ACTCAAAACAACCTTACACACACACACACACACACACACACTCACAAAGATA
TTTTGTTCACCATATGTATTGATGTGCCTGTTTTCAGTCAATCCACAGGAA
GGGCTAAGGAGAGTGACATCTGGGCTACATTAAAAGGACAGTCACATTG
CTCAAAGNACTCAAGTTTAGCCCGAGTACAGTAGCT
Genbank ID: Z24212
Description: H. sapiens (D18S477) DNA segment containing (CA) repeat; clone
Search for GDB entry

27. D18S979:

Database ID: GATA-P28080 (Also known as G08015, CHLC.GATA92C08, CHLC.GATA92C08.P28080)
Source: CHLC: genetically mapped polymorphic tetranucleotide repeats
Chromosome: Chr18

Primers:
Left = AGCTTGCAGATAGCCTGCTA
Right = TACGGTAGGTAGGTAGATAGATTCG
Product Length = 155
Review complete sequence:
CTCTACAGTCTCTNACCTTTGGACTCCAGGACTTTCACCAGCACCCTCAA
CATTCCCACTGGGTTCTCAGGACTTTATAGTTGTACTGAGCCATGCCACT
GGATCCTAGGGTCTCCAGCTTGCAGATAGCCTGCTATGGGACTTAATCT
TTGTAATAAGGTGAGTCAATTCTGCCAATAAACCTACTTTCATCTCTATCT
ATCTATCTATCTATCTATCTATCTATCTATATCTATCATCTATCTATCGAAT
CTATCTACCTACCTACCGTATTAGTTCTGTCTCTCTGGAGN

Genbank ID: G08015
Description: human STS CHLC.GATA92C08.P28080 clone GATA92C08.

28. WI-9340:

Database ID: UTR-05134 (Also known as G06102, D18S1034, 9340, X60221)
Source: WICGR: Primers derived from Genbank sequences
Chromosome: Chr18

Primers:
Left = TGAGAGAACGAAATCTCTATCGG
Right = AGGCAGCAAGTTTTTATAAAGGC
Product Length = 115
Review complete sequence:

ATGTATCTATCCCAATTGAGTCAGCTAGAAACAGTTGACTGACTAAATGG
AAACTAGTCTATTTGACAAAGTCTTTCTGTGTTGGTGTCTACTGAAGTTAT
AGTTTACCCTTCCTAAAAATGAAAAGTTTGTTTCATATAGTGAGAGAACGA
AATCTCTATCGGCCAGTCAGATGTTTCTCATCCTTCTTGCTCTGCCTTTG
AGTTGTTCCGTGATCATTCTGAATAAGCATTTGCCTTTATAAAAACTTGCT
GCCTGACTAAAGATTAACAGGTTATAGTTTAAATTTGTAATTAATTCTACC
ATCTTGCAATAAAGTGACAATTGAATG

Genbank ID: G06102
Description: WICGR: Random genome wide STSs
Search for GDB entry

29. D18S466:

Database ID: AFM094YE5 (Also known as 094ye5, Z23354, D18S466)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ACACTGTAGCAGAGGCTTGACC
Right = AGGCCAAGTTATGTGCCACC
Product Length = 214
Review complete sequence:
aaatgacactttaaggaggtaacactgtagcagaggcttgaccaccacccagttctcactagcactgagg
atgctctattggttgggttacccacacacgcatagacatgcacacacacagacacacagacacacacac
acacacacacacaccagatatagcattccaaaccatcaatatgctatgcaatactgcattaacaggtcatg
cctgtggtggcacataacttggcctagaaaatactggggacgtctgcattccctttttattatcgaattgacttact
tggcttctgagttttcctcagaagtaatacttcaatacctcttccatttctgccttgancattgtttggggtaccaag
tatagct

Genbank ID: Z23354
Description: H. sapiens (D18S466) DNA segment containing (CA) repeat; clone
Search for GDB entry

30. D18S1092:

Database ID: AFMA112WE9 (Also known as D18S1092, w5374, a112we9)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = CTCTCAAAGTAAGAGCGATGTTGTA
Right = CCGAAGTAGAAAATCTTGGCA
Product Length = 163
Review complete sequence:

agctctcaaagtaagagcgatgttgtaactgactgagttgttttgtgaantttttgnttttggagtcagtggagcat
gttattagatgtaaatttaaacacacacacacacacacacacacacacacacacacgagaagtaagtgccaag
attttctacttcggcgcctatatttctatatactgattttctgtatttcccagacttgaatatagattgtctttctgntttat
catagacaatctcataataanttaggcataataaggtaatgaggnttttctgggcttcttttcatcatccctgca
atttgagtctcntttatagntgaantcttctctgtaataacntcttgttttagct

Search for GDB entry

31. D18S61:

Database ID: AFM193YF8 (Also known as 193yf8, Z16834, D18S61)
Source: J Weissenbach, Genethon: genetically mapped polymorphic STSs
Chromosome: Chr18

Primers:
Left = ATTTCTAAGAGGACTCCCAAACT
Right = ATATTTTGAAACTCAGGAGCAT
Product Length = 174
Review complete sequence:
CGTCTTACCAAACCAACATAATATAGCAATGGNAACCAAAAATTTCTAAGA
GGACTCCCAAACTACATTCTTCTNCCTGAATTAAATACAGGCATTCAANA
NAAACANACACACACACACACACACACACACACACACACACACACGCACA
CCCTTCAAATCNTAGCATAAATTCCNCTTATATAAACATAACCATGCTCCT
GAGTTTCAAAATATTGGGTGGTTCGAAGTTCGAAGCAACAAATTTCCAGT
TAGTGTCTATTANTTGTTGGACAGCT

Genbank ID: Z16834
Description: H. sapiens (D18S61) DNA segment containing (CA) repeat; clone
Search for GDB entry

Markers (STRs) used in refining the candidate region.

Below the markers are shown that were used in family MAD31 to refine the candidate region. Most of these markers are already described above and will therefore only be mentioned to by their name. For the additional markers, the information is given here.

Data was already shown for: D18S68, D18S55, D18S969, D18S1113, D18S483, D18S465, D18S876, D18S477, D18S979, D18S466 and D18S61.

New data:

1. D18S51:

Other names: UT574, (D18S379)
Primer sequences:
UT574a GAGCCATGTTCATGCCACTG
UT574b CAAACCCGACTACCAGCAAC

DNA-sequence:
AATTGAGCNCAGGAGTTTAAGACCAGCCTGGGTAACACAGTGAGACCCC
TGTCTCTACAAAAAAATACAAAAATNAGTTGGGCATGGTGGCACGTGCCT
GTAGTCTCAGCTACTTGCAGGGCTGAGGCAGGAGGAGTTCTTGAGCCCA
GAAGGTTAAGGCTGCAGTGAGCCATGTTCATGCCACTGCACTTCACTCT
GAGTGACAAATTGAGACCTTGTCTCAGAAAGAAAGAAAGAAAGAAAGAAA
GAAAGAAAGAAAGAANGAAAGAAAGAAAGTAAGAAAAAGAGAGGGAAAG
AAAGAGAAANAGNAAANAAATAGTAGCAACTGTTATTGTAAGACATCTCC
ACACACCAGAGAAGTTAATTTTAATTTTAACATGTTAAGAACAGAGAGAAG
CCAACATGTCCACCTTAGGCTGACGGTTTGTTTATTTGTGTTGTTGCTGG
TAGTCGGGTTTGTTATTTTTAAAGTAGCTTATCCAATACTTCATTAACAAT
TTCAGTAAGTTATTTCATCTTTCAACATAAATACGNACAAGGATTTCTTCT
GGTCAAGACCAAACTAATATTAGTCCATAGTAGGAGCTAATACTATCACA
TTTACTAAGTATTCTATTTGCAATTTGACTGTAGCCCATAGCCTTTTGTCG
GCTAAAGTGAGCTTAATGCTGATCGACTCTAGAG

GENBANK ID: L18333

2. D18S346.

Other name: UT575

Primer Pairs:

Primer A: TGGAGGTTGCAATGAGCTG
Primer B: CATGCACACCTAATTGGCG

DNA sequence:
ACGAGGACAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCGTT
TNTACTAAAANTACAAAANTTGGTCGGGAGGCTGGGGCAGGNGACATGC

TTGACCCCAGGAGGTGGAGGTTGCAATGAGCTGAGATTGCACCACTGCA
CTNCAGCNTGG.....AAGAAAGAGAAAGGANAGNNAGGNAGNNANNAAAC
TACATNTGAAGTCAACACTAGTATTGGTGGGAGAGGAATTTTATGCTGCA
TTCCCCNACAACCACTAGATACGCCAATTAGGTGTGCATGGTCCATGCTA
T

GenBank ID: L26588

3. D18S817.

Other name: UT6365

Primer Pairs:

Primer A: GCAAAGCAGAAGTGAGCATG
Primer B: TAGGACTACAGGCGTGTGC

DNA Sequence:

CATATGGGTCCACAAGCAACCTCAGTCCTTGTCTCTTCAGAAGAAAGAAT
TCTACTGAGGGNCATAAGGCAGAAGGAGAGACCTAGGCAAGTTGCAAAG
CAGAAGTGAGCATGTATTAAAAAAGCTTTAGAACAGTAAGGAAAGGAAGAA
AAGAAAAGAAGGAAAGTTCAACTTGGAAGAGGGCCAAGCCGGCAACTTG
GCAGAAGGATTGCTTGAGCCCAGGAGTTAAGACCAGTCTGGGCAATATA
GTGAGACTCCATCTCTGCATACATACATACATACATACATACATACATACA
TACATACATATTGCAGGGTATGATGGCACACGCCTGTAGTCCTAGCTACT
CTGGAGGTTGAGATGGGAGGGTCACTGAGCCTGGGAANTTGAGGCTGC
NNTGAGCCATGATC

GenBank ID: L30552

Characterisation of YACs.

8 YACs were selected covering the candidate region and flanking the gap. These YACs were further characterised by determining the end-sequences by the Inverse-PCR protocol.
Selected YACs: 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752_g_8, 717_d_3, 745_d_2

New STSs based on end-sequences (unless indicated otherwise, the STSs were tested on a monochromosomal mapping pannel for identifying chimaerism of the YAC; if the STS revealed a hit not on chromosome 18q - chimaeric YAC- then it is indicated in the text below):

1. SV32L.

Derived from YAC 745_d_2 left arm end-sequence.

Primer A: GTTATTACAATGTCACCCTCATT
Primer B: ACATCTGTAAGAGCTTCACAAACA

DNA-sequence:

ATTCCTTNGTTATTACAATGTCACCCTCATTTAAAAAGTGGAAAGATAAAG
AGGAAGCAATCTATTTTTTTCCTTTTTTTCTGATAGCACTTGTTTGTGAAG
CTCTTACAGATGTTCTTAAGTAAAATCAACTCCTCCATTTTTTTTGTAGCA
ACTACACATATTTATCAATAATAGTTCACAAATACATTTTCAAATT

Amplified sequence length: 107 basepairs (bp)

This STS has no clear hit on the monochromosomal mapping pannel.

2. SV32R.

Derived from YAC 745_d_2 right arm end-sequence.

Primer A: ACGTTTCTCAATTGTTTAGTC
Primer B: TGTCTTGGCATTATTTTTAC

DNA sequence:

AGACAATGGGAGAAATTGCACTGCCCTGAGTCAGAAATCAGATCTGTTG
CCATACAGCTGCCGTTATGTGATCATTTGCAAGTCAACGTTTCTCAATTG
TTTAGTCATTTGTAAGACAAAAAGACTGGTTGGATTTCAGAGAATTTGGA
ATCCTCCTTCAGGTTTAACAAGCAATAAATGATACTCTTCAGTGTAAAAAT
AATGCCAAGACATNATTTGACTTTAAATTAAATCCAAACAAGATATC

Amplified sequence length: 127 bp

This STS has no clear hit on the monochromosomal mapping pannel.

3. SV11L.

Derived from YAC 766_f_12 left arm end-sequence.

Primer A: CTATGCTCTGATCTTTGTTACTTT
Primer B: ATTAACGGGAAAGAATGGTAT

DNA sequence:

GTCTTTATTTCATATAACTATGCTCTGATCTTTGTTACTTTCTCCTTTTAAC
TCAGTTTAAGCTTTATTCTTATTTTCCAGCTGCTGAAGGTATATAGTTAGG
TTGTTTATTGGATACCATTCTTTCCCGTTAATGTCAGTGGTTACTGCTATC
AATGTAGCAGTTA

Amplified sequence length: 118 bp

This STS has a hit with chromosome 18 and must ba located between CHLC.GATA-p6051 and D18S968.

4. SV11R.

Derived from YAC 766_f_12 right arm end-sequence.

Primer A: AAGGTATATTATTTGTGTCG
Primer B: AAACTTTTCTTAACCTCATA

DNA sequence:

ATAAGGTATATTATTTGTGTCGTGAGTTAAGAAATCATTAATAACTATTTT
CAGAATGACAAATGTCATTATATGTTGTAAAAAAGATAAATACGTGAAATT
ATGAGGTTAAGAAAAGTTTA

Amplified sequence length: 119 bp.

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01.

5. SV34L.

Derived from YAC 717_d_3 left arm end-sequence.

Primer A: TCTACACATATGGGAAAGCAGGAA
Primer B: GCTGGTGGTTTTGGAGGTAGG

ACATAAAATGTCGCTCAAAAACAATTATGTGTGTCTACACATATGGGAAA
GCAGGAAACAAATTTGTTTACAACATACATTACTTTTGTTTTTTAGGCAAG
ATAAAATNTCCTACCTCCAAAACCACCAGCACNGTCCGCAATAACTATAC
ATC

Amplified sequence length: 98 bp

This STS has a hit with chromosome 18.

6. SV34R.

Derived from YAC 717_d_3 right arm end-sequence.

Primer A: ATAAGAGACCAGAATGTGATA
Primer B: TCTTTGGAGGAGGGTAGTC

DNA-sequence:

AATATCATTCTTCACCCACGTTATACATAAGAGACCAGAATGTGATATTGT
CATCTCACATGGAAAAATCTGCTGTGATCAGTTCCTGAAGCTTGCTGTGA
TCCTCCCTTAGGAAAGTAGAAAAATCTTTTTGAAACACTTTATTCTACAAT
CAATGAAAATTAGGTGAAGCTACAGAAGCCAGAAATTACTCTAAGATTAG
ACAATTATTTAAGANGACCAATTGTCTTTGGTCTTCTTCTGAAGGGTCTG
ACTACCCTCCTCCAAAGAATTCACTGGCCGTCGTTTTACAACGTCNTGA

Amplified sequence length: 244 bp

This STS has a hit with chromosome 1, therefore YAC 717_d_3 is chimaeric

7. SV25L.

Derived from YAC 731_c_7 left arm end-sequence.

Primer A: AAATCTCTTAAGCTCATGCTAGTG
Primer B: CCTGCCTACCAGCCTGTC

DNA sequence:

AGTGGAGAGATAGAAAGAGAGGAAGATTTTTTTTTTTTAAATCTCTTAAGCT
CATGCTAGTGTAGGTGCTGGCAGGTCTGAACACTCTGTAGGACAGGCTG
GTAGGCAGGAA

Amplified sequence length: 72 bp

This STS has no clear hits on the monochromosomal mapping pannel.

8. SV25R.

Derived from YAC 731_c_7 right arm end-sequence.

Primer A: TGGGGTGCGCTGTGTTGT
Primer B: GAGATTTCATGCATTCCTGTAAGA

DNA-sequence:

GGAGGGTGTTNTCACANAAGTCTGGGGTGCGCTGTGTTGTTCATTGTAA
AAACCCTTTGGANCATCTGGGAATGTGCTGCCCCACATGTCCAGGTAAC
GTTCTCAGGAAGGGGAGGCTGGAAATCTCTGTGTGTTCTTACAGGAATG
CATGAAATCTCCCANCCCCTCTTGTTGGAAATTTCCCTCACTTT

Amplified sequence length: 136 bp

This STS has a hit with chromosome 7; therefore YAC 731_c_7 is chimaeric

9. SV31L.

Derived from YAC 752_g_8 left arm end-sequence.

Primer A: GAGGCACAGCTTACCAGTTCA
Primer B: ATTCATTTTCTCATTTTATCC

DNA-sequence:

CTTCTCNATGANTGGACAAATGTCATTGGGTCAGCATGAGGCACAGCTT
ACCAGTTCAGATTCCAGTAGCTGAGGAACAAATCTTAACTCCAAAAATAA
GTAATTGCGTCACTTTGGAGGAATTATTTGACCTTTTCATAACTTTGACAT
CACAACAATGAGGGTGAAGTTAGTAAAATAAATGATTATTATGAGGATAA
AATGAGAAAATGAATTNAGTGCTTAAGACAATGCTTGGTAACTAGTTAAN
CCG

Amplified sequence length: 178 bp

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01.

10. SV31R.

Derived from YAC 752_g_8 right arm end-sequence.

Primer A: CAAGATTATGCCTCAACT
Primer B: TAAGCTCATAATCTCTGGA

DNA sequence:

AAACTTTAACCAATTTAAACTCCCTAACAGTTCTATAAAATAAGCAAGATT
ATGCCTCAACTTTATGGATAAAGAAATGGAGGCATTAAGAGATAACTAAC
TTGCCCAAGGCCACACAAGTGACTGAGTAAGAATTGCAAAGCCAATGAG
TCTGGCTCCAGAGATTATGAGCTTAATCACCACACTGTGCCACCTCCTGT
GTTTCCTGG

Amplified sequence length: 131 bp

This STS has no clear hits on the monochromosomal mapping pannel and gives no information concerning the chimaerity of the YAC.

11. SV10L.

Derived from YAC 942_c_3 left arm end-sequence.

Primer A: TCACTTGGTTGGTTAACATTACT
Primer B: TAGAAAAACAGTTGCATTTGATAT

DNA-sequence:

GGTNTTTCACTTGGTTGGTTAACATTACTTCTAAGTTTTTTATTGTTTTTTA
TGCTATTGCTAATGGGATTGCTTTCTTAATTTATTTTTTCCAATAGCTTGT
TGTTAGTTTATATCAAATGCAACTGTTTTTCTATGCAAATTATGTTTCCT

Amplified sequence length: 130 bp

This STS has a hit with chromosome 18 and must be located between CHLC.GATA-p6051 and D18S968

12. SV10R.

Derived from YAC 942_c_3 right arm end-sequence.

Primer A: AACCCAAGGGAGCACAACTG
Primer B: GGCAATAGGCTTTCCAACAT

DNA sequence:
TTGGTGGTGCCCTAGGTTTGGCAATTATAAATAAAGCTGCTACAAACATT
CATGTGCAGGTCTCCGTGTGGACATAATTTTCCAGTTCATTTGGGTAAAA
CCCAAGGGAGCACAACTGTTGGATCCTATNATAAAAATATNTCTCGTTTC
ATTTAAAAAACCTGGGAAACTATCTNCCCACAGTGGCTGTCCCTTTTTGT
ATCCCCACCAACAATGTTGGAAAGCCTATTGCCANCAT

Amplified sequence length: 135 bp

This STS has a hit with chromosome 18 and must be located between D18S876 and GCT3G01

13. SV6L.

Derived from YAC 961_h_9 left arm end-sequence.

No primer was made, because this sequence is identical to a known STR marker D18S42, which is indeed mapped to this region.

Primer A:
Primer B:

DNA sequence:

CATGNCTCACAGTGTTCTGAGGCTGCTCTGGACATGCAATCTTGCATGC
TTTTGTCATGACAGGTCTTAAANAGTTTATCAGCTTNCTCAAATAGCTGAA
TGACANAACACTGGATTTTTGTTCAAATANCCTATCAACTTGGCNTCTGT
GTTGCGGTTGTCACTTGGTAACAAAATAAGTC

Amplified sequence length:

SV6L recognises D18S42 which must be therefore located between WI-7336 and WI-8145

14. SV6R.

Derived from YAC 961_h_9 right arm end-sequence.

Primer A: TTGTGGAATGGCTAAGT
Primer B: GAAAGTATCAAGGCAGTG

DNA sequence:

TAATTGACAAATAAAAATTGTATATTTTNCATATTTAACATGTTATGCTAAC
ATATATATGGATTGTGGAATGGCTAAGTCAGAAATTCTTTTACATTCATAT
TTCCATATTATTTACTTTNNGCTTTAAAAAATATGTAAATGANAATACTTAT
TTTTTTCAGTGTCACTGCCTTGATACTTTCACATTTNNGTTACATATTATTT
CCCTTNCATCTAACAAATATATATTGAGTTTCTATAATGTGTCTGACACTG
A

Amplified sequence length: 122 bp

SV6R amplifies a segment on chromosome 18. This segment must be located between WI-2620 and WI-4211

15. SV26L.

Derived from YAC 907_e_1 left arm end-sequence.

Primer A: TATTTGGTTTGTTTGCTGAGGT
Primer B: CAAGAAGGATGGATACAAACAAG

DNA sequence:

TGGTCACTGGTGCCTTATTTGGTTTGTTTGCTGAGGTCATATTTCCTGTG
GCCTTCATGCTTGATTTGTTGGAGTCTAGCCATGTAAAANTCTGTTGGAG
TCTAGGCATTTAAAAAATAGGTATTTATTGTAATCTTTGCCATTTGCTTGT
TTGTATCCATCCTTCTTGGGAAGGCTTTACAGGCATTCAAAAGG

Amplified sequence length: 154 bp

This STS has a hit with chromosome 13; therefore YAC 907_e_1 is chimaeric.

16. SV26R.

Derived from YAC 907_e_1 right arm end-sequence.

Primer A: CGCTATGCATGGATTTA
Primer B: GCTGAATTTAGGATGTAA

DNA sequence:

CGCTATGCATGGATTTAAACTGAGTGTAGTGCACTCACTATGTTGCAGTC
TCTTATTCTAGGTTCCTAATATTTACATCCTAAATTCAGCT

Amplified sequence length: 90 bp

no clear hits on monochromosomal mapping pannel: no information concerning chaemerity at this side of the YAC

Testing of 3 end-sequences flanking the gap in additional YACs: STS-markers WI-4211, D18S876 and GCT3G01 are also shown in order to identify YACs on opposite sides of the gap more clearly in table 3 below.

	STSs					
YACs	WI-4211	D18S876	SV31L	SV11R	SV10R	GCT3G01
940_b_1	+	+	+	-	-	-
766_f_12	+	+	+	+	-	-
846_a_5	+	- ?	+	+	-	-
752_g_8	+	+	+	+	-	-
745_d_2	+	+	+	+	-	-
961_c_1	+	+	-	-	-	-
942_c_3	+	+	+	+	+	-
717_d_3	-	-	+	+	- ?	+
972_e_11	-	-	-	-	-	+
940_h_10	-	-	-	-	+	+
821_e_7	-	-	-	-	+	+
731_c_7	-	-	-	-	-	+
889_c_4	-	-	-	-	+	+
907_e_1	-	-	-	+	+	+

- +: positive hit / -: no hit / ?: 2 instances were observed in which a positive hit was expected (on the assumed order of the markers) but not observed. The reasons for this are not clear.

YAC 745_d_2 was excluded from further analysis since there was no clear hit with chromosome 18. Of the remaining 7 from a monochromosomal mapping panel it was determined that 3 were chimeric and 4 non-chimeric as shown in Table 4 below.

TABLE 4

YAC	chimaeric	chromosome
961_h_9 (6)	no	
942_c_3 (10)	no	
766_f_12 (11)	no	
731_c_7 (25)	yes	chromosome 7
907_e_1 (26)	yes	chromosome 13
752_g_8 (31)	no	
717_d_3 (34)	yes	chromosome 1

For the non-chimeric YACs the STS based on the end-sedquence flanking the gap (10R, 11R, 31L) was tested on 14 YACs flanking the gap. Overlaps between YACs on opposite sides of the gap were demonstrated: e.g. the "11R" end-sequence (766.f.12) detects YAC 766.f.12 and YAC 907.e.1.
YACs were then selected comprising the minimum tiling path:

TABLE 5

YAC	size	chimaerity
961_h_9	1180 kb	not chimaeric
766_f_12	1620 kb	not chimaeric
907_e_1	1690 kb	chimaeric (chr. 13)

These three YACs are stable as determined by PFGE and their sizes roughly correspond to the published sizes. These YACs were transferred to other host-yeast strains for restriction mapping.

Experimental 2

Construction of fragmentation vector:

A 4.5kb ECORI/SalI fragment of pBLC8.1 (Lewis *et al*, 1992) carrying a lysine-2 and a telomere sequence was directionally cloned into GEM3zf(-) digested with ECORI/SalI. Subsequently, an End Rescue Site was ligated into the EcoRI site. Hereto, two oligonucleotides (strand 1: 5'-TTCGGATCCGGTACCATCGAT-3' AND STRAND 2: 3'-GCCTAGGCCATGGTAGCTATT-5') were ligated into a partial (dATP) filled ECORI site, generating the vector pDF1. Triplet repeat containing fragmentation vectors were constructed by cloning of a 21bp and a 30bp CAG/CTG adapter into the Klenow-filled PstI site of pDF1. Trasformation and selection resulted in a $(CAG)_7$ and a $(CTG)_{10}$ fragmentation vector with the orientation of the repeat sequence 5' to 3' relative to the telomere.

Yeast transformation:

Linearised (digested with SalI) vector was used to transform YAC clones 961_h_9, 766_f_12 or 907_e_1 using the LiAc method. After transformation the YAC clones were plated onto SDLys$^-$ plates to select for the presence of the fragmentatio vector. After 2-3 days colonies were replica plated onto SDLys$^-$-Trp$^-$-Ura$^-$ and SDLys$^-$-Trp$^-$-Ura$^+$ plates. Colonies growing on the SDLys$^-$-Trp$^-$-Ura$^+$ plates but not on the SDLys$^-$-Trp$^-$-Ura$^-$ plates contained the fragmented YACs.

Analysis of fragmented YACs:

Yeast DNA isolated from clones with the correct

phenotype was analysed by Pulsed Field Electrophoresis (PFGE), followed by blotting and hybridisation with the Lys-2 gene and the sizes of the fragmented YACs were estimated by comparison with DNA standards of known length.

End Rescue:

Fragmented YACs characterised by a size common to other fragmented YACs, indicative of the presence of a major CAG or CTG triplet repeat, were digested with one of the enzymes from the End Rescue site, ligated and used to transform E. Coli. After growth of the transformed bacteria the plasmid DNA was isolated and the ends of the fragmented YACs, corresponding to one of the sequences flanking the isolated trinucleotide repeats, were sequenced.

Sequencing revealed that fragmented YACs of an equal length were all fragmented at the same site. A BLAST Search of the GenBank database was performed with the identified sequences to identify homology with known sequences. The complete sequence spanning the CAG or CTG repeats of the fragmented YACs was obtained by Cosmid Sequencing, employing sequence specific primers and splice primers, as previously described (Fuentes *et al*. 1992 Hum.Genet. **101**: 346-350) or by using the "genome walker" kit (Clontech Laboratories, Palo Alto, USA) and described in Siebert *et al*. Nucleic Acid Res (1995) **23**(6): 1087-1088 and Siebert *et al*. (1995) CLONTECHniques **X**(II): 1-3.

Results:

A YAC 961_h_9 clone was transformed with the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. The CTG vector

did not reveal the presence of any CTG repeat. Analysis of twelve $(CAG)_7$ fragmented YACs showed that five of these had the same size of approximately 100kb. End Rescue was performed with ECORI and sequencing of three of these fragments revealed that they all shared the terminal sequence shown in italics in Figure 15a. A BLAST search of the Genbank database with this sequence indicated the presence of a sequence homology with the CAP2 gene (GenbBank accession number: L40377). The sequence spanning the CAG repeat shown in Figure 15a was obtained by both cosmid sequencing and genome walker sequencing. The sequence was mapped between markers D18S68 and WI-3170 by STS content mapping.

A YAC 766-f-12 was fragmented using the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. Again the $(CTG)_{10}$ vector did not reveal the presence of any CTG repeat. Analysis of twenty $(CAG)_7$ fragmented YACs showed the presence of two groups of fragments with the same size: five of approximatively 650kb and two of approximatively 50kb.

End Rescue was performed using ECORI on four of the fragmented YACs of 650kb. Sequencing confirmed that they all shared identical 3' terminals, characterised by the sequence shown in italics in Figure 16a. A Blast Search showed homology of this sequence with the Alu repeat sequence family. The sequence spanning the CAG repeat shown in Figure 16a was obtained by cosmid sequencing. The sequence was mapped between markers WI-2620 and WI-4211 by STS content mapping on the YAC contig map.

End Rescue was also performed on the two fragments of 50kb. Sequencing revealed the sequence shown in italics in figure 17a. A Blast Search revealed no

sequence homology with any known sequence. Cosmid sequencing allowed to identify the complete sequence spanning the CAG repeats, shown in figure 17a. The sequence was mapped between markers D18S968 and D18S875 by STS content mapping on the YAC contig map.

A YAC 907-e-1 clone was transformed with the $(CAG)_7$ or $(CTG)_{10}$ fragmentation vector. The $(CAG)_7$ vector did not reveal the presence of any CAG repeat. Analysis of twenty-six $(CTG)_{10}$ fragmented YACs revealed that twenty-one of them had the same size of approximatively 900kb. End Rescue was performed with KpnI on three fragmented YACS of this size. Sequencing revealed the nucleotide sequence shown in italics in Figure 18a. A Blast Search indicated the presence of an homology of this sequence with the GCT3G0I marker (GenBank accession number: G09484). The sequence spanning the CTG repeat was obtained from the GenBank Database. The sequence was mapped between markers 10R and WI-528.

CLAIMS:

1. Use of an 8.9 cM region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979 or a fragment thereof for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder.

2. Use of a YAC clone comprising a portion of human chromosome 18q disposed between polymorphic markers D18S60 and D18S61 for identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder.

3. The use as claimed in claim 2 wherein said portion comprises the region of chromosome 18q between polymorphic markers D18S68 and D18S979 or a fragment of said region.

4. The use as claimed in claim 2 or 3 wherein said YAC clone is 961_h_9, 942_c_3, 766_f_12, 731_c_7, 907_e_1, 752-g-8 or 717_d_3.

5. The use as claimed in claim 4 wherein said YAC clone is 961_h_9, 766_f_12 or 907_e_1.

6. The use as claimed in any preceding claim wherein said mood disorder or related disorder is selected from the Diagnostic and Statistical Manual of Mental Disorders, version 4 (DSM-IV) taxonomy and includes mood disorders (296.XX, 300.4, 311, 301, 13, 295.70), schizophrenia and related disorders (295, 297.1, 298.9, 297.3, 298.9), anxiety disorders

(300.XX, 309.81, 308.3), adjustment disorders (309, XX) and personality disorders (codes 301. XX).

7. A method of identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which comprises detecting nucleotide triplet repeats in a region of human chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

8. A method of identifying at least one human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which comprises fragmentation of a YAC clone as defined in any one of claims 2 to 4 and detection of nucleotide triplet repeats.

9. A method as claimed in claim 7 or 8 wherein said repeated triplet is CAG or CTG.

10. A method as claimed in claim 9 wherein said repeated triplet is detected by means of a probe comprising at least 5 CTG and/or CAG repeats.

11. A method of identifying at least one human gene including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder wherein said gene is present in the DNA comprised in the YAC clones as defined in any one of claims 2 to 5, which method comprises the step of detecting an expression product of said gene with an antibody capable of recognising a protein with an amino acid sequence comprising a string of at least 8 continuous glutamine residues.

12. A method as claimed in claim 11 wherein said DNA forms part of a human cDNA expression library.

13. A method as claimed in claim 11 or claim 12 wherein said antibody is mAB 1C2.

14. A method of preparing a contig map of YAC clones of the region of human chromosome 18q between polymorphic markers D18S60 and D18S61 which comprises the steps of:

(a) subcloning the YAC clones according to any one of claims 2 to 5 into exon trap vectors;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other known sequence tagged sequence from the YAC contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to detect overlaps among the cosmid vectors, and

(c) constructing a cosmid contig map of a YAC clone of said region.

15. A method of identifying at least one human gene or any mutated or polymorphic variant thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) transfecting mammalian cells with DNA sequences cloned into an exon trap vector as prepared in claim 14;

(b) culturing said mammalian cells in an appropriate medium;

(c) isolating RNA transcripts expressed from an SV40 promoter;

(d) preparing cDNA from said RNA transcripts;

(e) identifying splicing events involving exons of the DNA subcloned into said exon trap vector in accordance with claim 14 to elucidate positions of coding regions in said subcloned DNA;

(f) detecting differences between said coding regions and equivalent regions in the DNA of an individual afflicted with said mood disorder or related disorder; and

(g) identifying said gene or mutated or polymorphic variants thereof which is associated with said mood disorder or related disorder.

16. A method of identifying at least one human gene or mutated or polymorphic variants thereof which is associated with a mood disorder or related disorder which comprises the steps of:

(a) subcloning the YAC clones according to any one of claims 2 to 5 into a cosmid, BAC, PAC or other vector;

(b) using the nucleotide sequences shown in any one of Figures 1 to 11 or any other known sequence tagged sequence from the YAC contig described herein, or part thereof consisting of not less than 14 contiguous bases or the complement thereof, to defect overlaps amongst the subclones and construct a map

thereof;

(c) identifying the position of genes within the subcloned DNA by one or more of CpG island identification, zoo-blotting, hybridization of said subcloned DNA to a cDNA library or a Northern blot of mRNA from a panel of culture cell lines;

(d) detecting differences between said genes and equivalent regions of the DNA of an individual afflicted with a mood disorder or related disorder; and

(e) identifying said gene which, if defective, is associated with said mood disorder or related disorder.

17. An isolated human gene, including mutated or polymorphic variants thereof, which is associated with a mood disorder or related disorder which is obtainable by the method according to any of claims 7 to 13, 15 or 16.

18. A human protein which, if defective, is associated with a mood disorder or related disorder which is the expression product of the gene according to claim 17.

19. A cDNA encoding the protein of claim 18 which is obtainable by the method of any one of claims 7 to 13 , 15 or 16.

20. Use of a probe of at least 14 contiguous nucleotides of the cDNA of claim 19 or the complement thereof in a method for detection in a patient of a

pathological mutation or genetic variation associated with a mood disorder or related disorder which method comprises hybridizing said probe with a sample from said patient and from a control individual.

21. A nucleic acid molecule which comprises a sequence of nucleotides as shown in any one of Figures 15a, 16a, 17a or 18a.

22. A nucleic acid molecule which comprises a sequence of nucleotides which differ from a sequence of nucleotides as shown in any one of Figures 15a, 16a, 17a or 18a only in the extent of trinucleotide repeats.

23. A protein encoded by a nucleic acid molecule as claimed in claim 21.

24. A protein encoded by a nucleic acid molecule as claimed in claim 22.

25. A method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises analysing a sample of DNA from that individual for the presence of a DNA polymorphism associated with a mood disorder or related disorder in a region of chromosome 18q disposed between polymorphic markers D18S68 and D18S979.

26. A method as in claims 25 wherein said DNA polymorphism is a trinucleotide repeat expansion.

27. A method as in claim 26 wherein said trinucleotide repeat expansion is comprised in a

sequence of nucleotides that differ from the sequence of nucleotides shown in any one of Figures 15a, 16a, 17a or 18a only in said trinucleotide repeat expansion.

28. A method as in claim 26 or 27 which comprises the steps of:

a) obtaining a DNA sample from said individual;

b) providing primers suitable for the amplification of a nucleotide sequence comprised in the sequence shown in any one of Figures 15a, 16a, 17a or 18a said primers flanking the trinucleotide repeats comprised in said sequence;

c) applying said primers to the said DNA sample and carrying out an amplification reaction;

d) carrying out the same amplification reaction on a DNA sample from a control individual; and

e) comparing the results of the amplification reaction for the said individual and for the said control individual;

wherein the presence of an amplified fragment from said individual which is bigger in size from that of said control individual is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

29. A method as in claim 28 wherein said

nucleotide sequence to be amplified is comprised in the sequence shown in Figure 15a and said primers have the sequences shown in Figure 15b.

30. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 16a and said primers have the sequences shown in Figure 16b.

31. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 17a and said primers have the sequences shown in Figure 17b.

32. A method as in claim 28 wherein said nucleotide sequence to be amplified is comprised in the sequence shown in Figure 18a and said primers have the sequences shown in Figure 18b.

33. A method of determining the susceptibility of an individual to a mood disorder or related disorder which method comprises the steps of :

a) obtaining a protein sample from said individual; and

b) detecting the presence of the protein of claim 24;

wherein the presence of said protein is an indication of the presence of a susceptibility to a mood disorder or related disorder of said individual.

34. A method as in claim 33 wherein said protein is detected with an antibody which is capable of

recognising a string of at least 8 continuous glutamines.

35. A method as in claim 34 wherein said antibody is mAB 1C2.

36. A nucleic acid as claimed in claim 21 for use as a medicament in the treatment of a mood disorder or related disorder.

37. A protein as claimed in claim 23 for use as a medicament in the treatment of a mood disorder or related disorder.

38. A pharmaceutical composition which comprises a nucleic acid as claimed in claim 21 and a pharmaceutically acceptable carrier.

39. A pharmaceutical composition which comprises a protein as claimed in claim 23 and a pharmaceutically acceptable carrier.

40. An expression vector which comprises a sequence of nucleotides as claimed in claims 21 or 22.

41. A reporter plasmid which comprises the promoter region of a nucleic acid molecule as claimed in claim 21 or 22 positioned upstream of a reporter gene which encodes a reporter molecule so that expression of said reporter gene is controlled by said promoter region.

42. A cell line transfected with the expression vector of claim 40.

43. An eukaryotic cell or multicellular tissue or organism comprising a transgene encoding a protein as claimed in claims 23 or 24.

44. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell as claimed in claim 42 with said compound;

b) detecting and/or quantitatively evaluating the presence of any mRNA transcript corresponding to a nucleic acid as claimed in claim 21 or 22; and

c) comparing the level of transcription of said nucleic acid with the level of transcription of the same nucleic acid in a cell as claimed in claim 42 not exposed to said compound;

45. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell as claimed in claim 42 with said compound;

b) detecting and/or quantitatively evaluating the expression of a protein as claimed in claims 23 or 24 and

c) comparing the level of expression of said

protein with that of the same protein in a cell not exposed to said compound.

46. A method for determining if a compound is an enhancer or inhibitor of expression of a gene associated with a mood disorder or related disorder which comprises the steps of:

a) contacting a cell transfected with a reporter plasmid as claimed in claim 41 with said compound;

b) detecting or quantitatively evaluating the amount of reporter molecule expressed; and

c) comparing said amount with the amount of expression of said reporter molecule in a cell comprising said reporter plasmid and not exposed to said compound.

47. A compound identified as an enhancer or an inhibitor of the expression of a gene associated with a mood disorder or related disorder by a method as claimed in claims 44 to 46.

FIG. 1.

GTCTTTATTTCATATAACTATGCTCTGATCTTTGTTACTTTCTCCTTTTAAC
TCAGTTTAAGCTTTATTCTTATTTTCCAGCTGCTGAAGGTATATAGTTAGG
TTGTTTATTGGATACCATTCTTTCCCGTTAATGTCAGTGGTTACTGCTATC
AATGTAGCAGTTA

FIG. 2.

ATAAGGTATATTATTTGTGTCGTGAGTTAAGAAATCATTAATAACTATTTT
CAGAATGACAAATGTCATTATATGTTGTAAAAAAGATAAATACGTGAAATT
ATGAGGTTAAGAAAAGTTTA

FIG. 3.

ACATAAAATGTCGCTCAAAAACAATTATGTGTGTCTACACATATGGGAAA
GCAGGAAACAAATTTGTTTACAACATACATTACTTTTGTTTTTTAGGCAAG
ATAAAATNTCCTACCTCCAAAACCACCAGCACNGTCCGCAATAACTATAC
ATC

FIG. 4.

AATATCATTCTTCACCCACGTTATACATAAGAGACCAGAATGTGATATTGT
CATCTCACATGGAAAAATCTGCTGTGATCAGTTCCTGAAGCTTGCTGTGA
TCCTCCCTTAGGAAAGTAGAAAAATCTTTTTGAAACACTTTATTCTACAAT
CAATGAAAATTAGGTGAAGCTACAGAAGCCAGAAATTACTCTAAGATTAG
ACAATTATTTAAGANGACCAATTGTCTTTGGTCTTCTTCTGAAGGGTCTG
ACTACCCTCCTCCAAAGAATTCACTGGCCGTCGTTTTACAACGTCNTGA

FIG. 5.

GGAGGGTGTTNTCACANAAGTCTGGGGTGCGCTGTGTTGTTCATTGTAA
AAACCCTTTGGANCATCTGGGAATGTGCTGCCCCACATGTCCAGGTAAC
GTTCTCAGGAAGGGGAGGCTGGAAATCTCTGTGTGTTCTTACAGGAATG
CATGAAATCTCCCANCCCCTCTTGTTGGAAATTTCCCTCACTTT

FIG. 6.

CTTCTCNATGANTGGACAAATGTCATTGGGTCAGCATGAGGCACAGCTT
ACCAGTTCAGATTCCAGTAGCTGAGGAACAAATCTTAACTCCAAAAATAA
GTAATTGCGTCACTTTGGAGGAATTATTTGACCTTTTCATAACTTTGACAT
CACAACAATGAGGGTGAAGTTAGTAAAATAAATGATTATTATGAGGATAA
AATGAGAAAATGAATTNAGTGCTTAAGACAATGCTTGGTAACTAGTTAAN
CCG

FIG. 7.

GGTNTTTCACTTGGTTGGTTAACATTACTTCTAAGTTTTTTATTGTTTTTTA
TGCTATTGCTAATGGGATTGCTTTCTTAATTTATTTTTTCCAATAGCTTGT
TGTTAGTTTATATCAAATGCAACTGTTTTTCTATGCAAATTATGTTTCCT

FIG. 8.

TTGGTGGTGCCCTAGGTTTGGCAATTATAAATAAAGCTGCTACAAACATT
CATGTGCAGGTCTCCGTGTGGACATAATTTTCCAGTTCATTTGGGTAAAA
CCCAAGGGAGCACAACTGTTGGATCCTATNATAAAAATATNTCTCGTTTC
ATTTAAAAAACCTGGGAAACTATCTNCCCACAGTGGCTGTCCCTTTTTGT
ATCCCCACCAACAATGTTGGAAAGCCTATTGCCANCAT

FIG. 9.

CATGNCTCACAGTGTTCTGAGGCTGCTCTGGACATGCAATCTTGCATGC
TTTTGTCATGACAGGTCTTAAANAGTTTATCAGCTTNCTCAAATAGCTGAA
TGACANAACACTGGATTTTTGTTCAAATANCCTATCAACTTGGCNTCTGT
GTTGCGGTTGTCACTTGGTAACAAAATAAGTC

FIG. 10.

TAATTGACAAATAAAAATTGTATATTTTNCATATTTAACATGTTATGCTAAC
ATATATATGGATTGTGGAATGGCTAAGTCAGAAATTCTTTTACATTCATAT
TTCCATATTATTTACTTTNNGCTTTAAAAAATATGTAAATGANAATACTTAT
TTTTTTCAGTGTCACTGCCTTGATACTTTCACATTTNNGTTACATATTATTT
CCCTTNCATCTAACAAATATATATTGAGTTTCTATAATGTGTCTGACACTG
A

FIG. 11.

TGGTCACTGGTGCCTTATTTGGTTTGTTTGCTGAGGTCATATTTCCTGTG
GCCTTCATGCTTGATTTGTTGGAGTCTAGCCATGTAAAANTCTGTTGGAG
TCTAGGCATTTAAAAAATAGGTATTTATTGTAATCTTTGCCATTTGCTTGT
TTGTATCCATCCTTCTTGGGAAGGCTTTACAGGCATTCAAAAGG

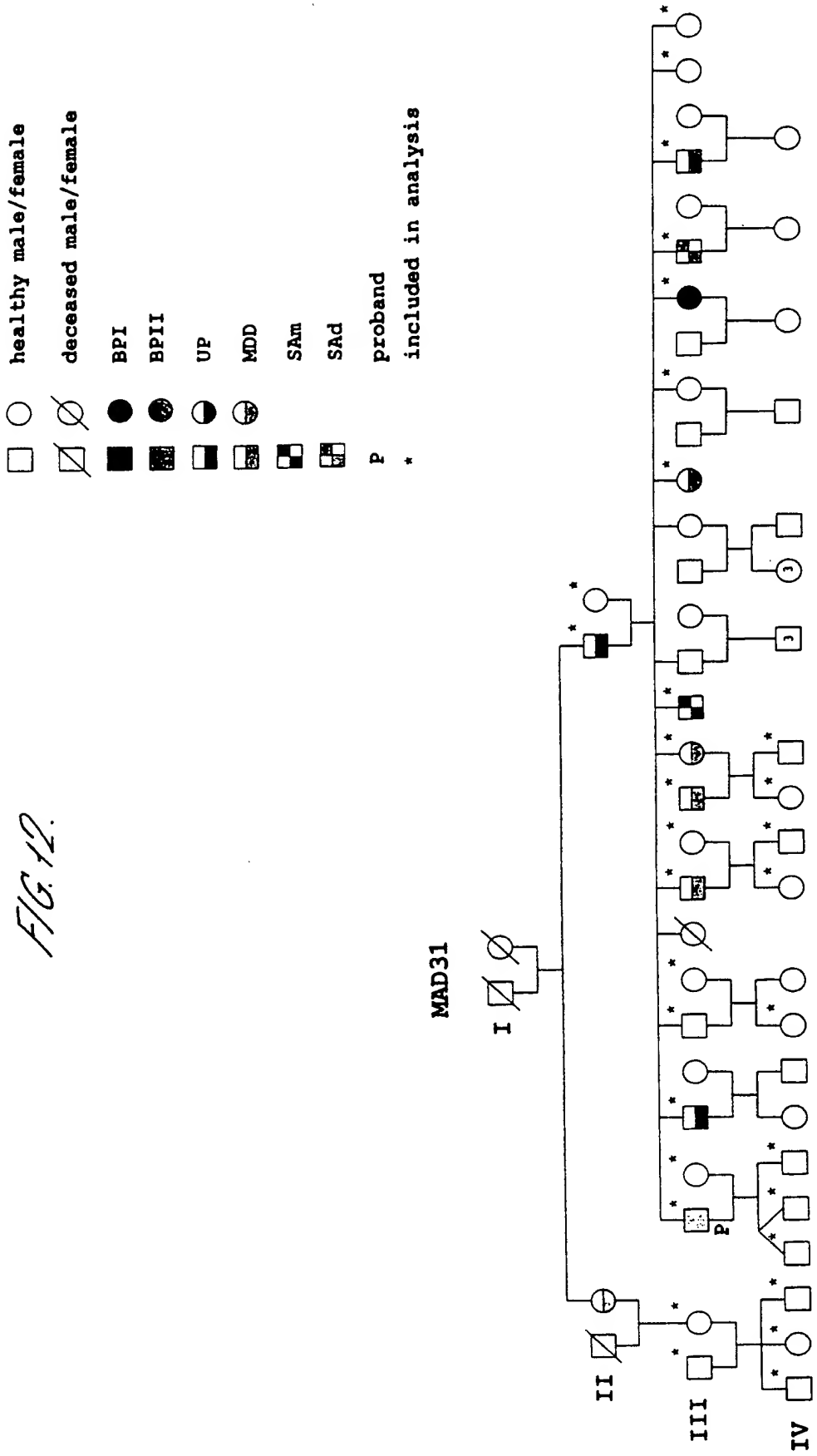
healthy male/female
deceased male/female
BPI
BPII
UP
MDD
SAm
SAd
proband
included in analysis
MAD31
I
II
III
IV
P

FIG. 12.

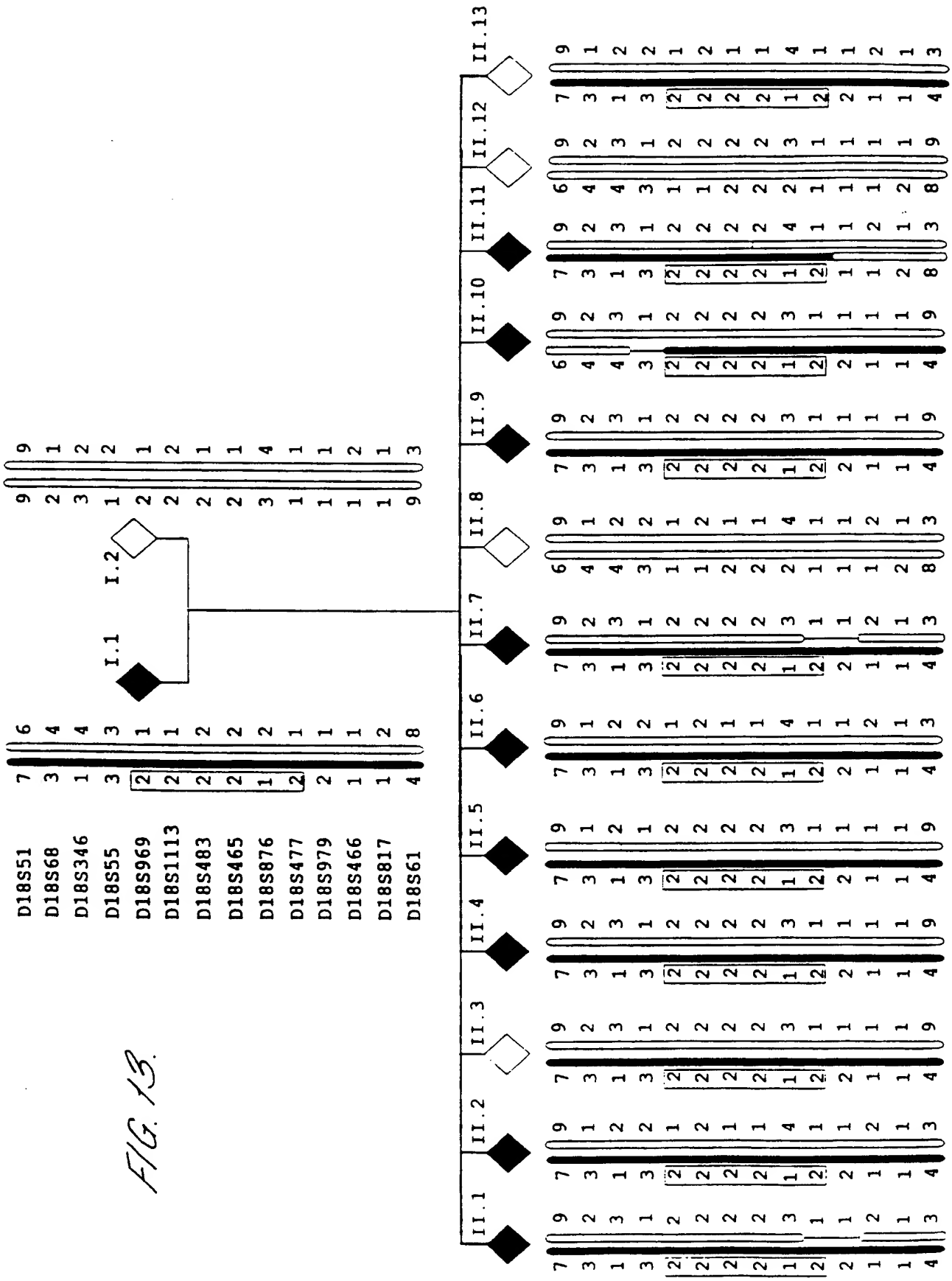
D18S51
D18S68
D18S346
D18S55
D18S969
D18S1113
D18S483
D18S465
D18S876
D18S477
D18S979
D18S466
D18S817
D18S61
I.1
I.2
II.1
II.2
II.3
II.4
II.5
II.6
II.7
II.8
II.9
II.10
II.11
II.12
II.13

FIG. 13.

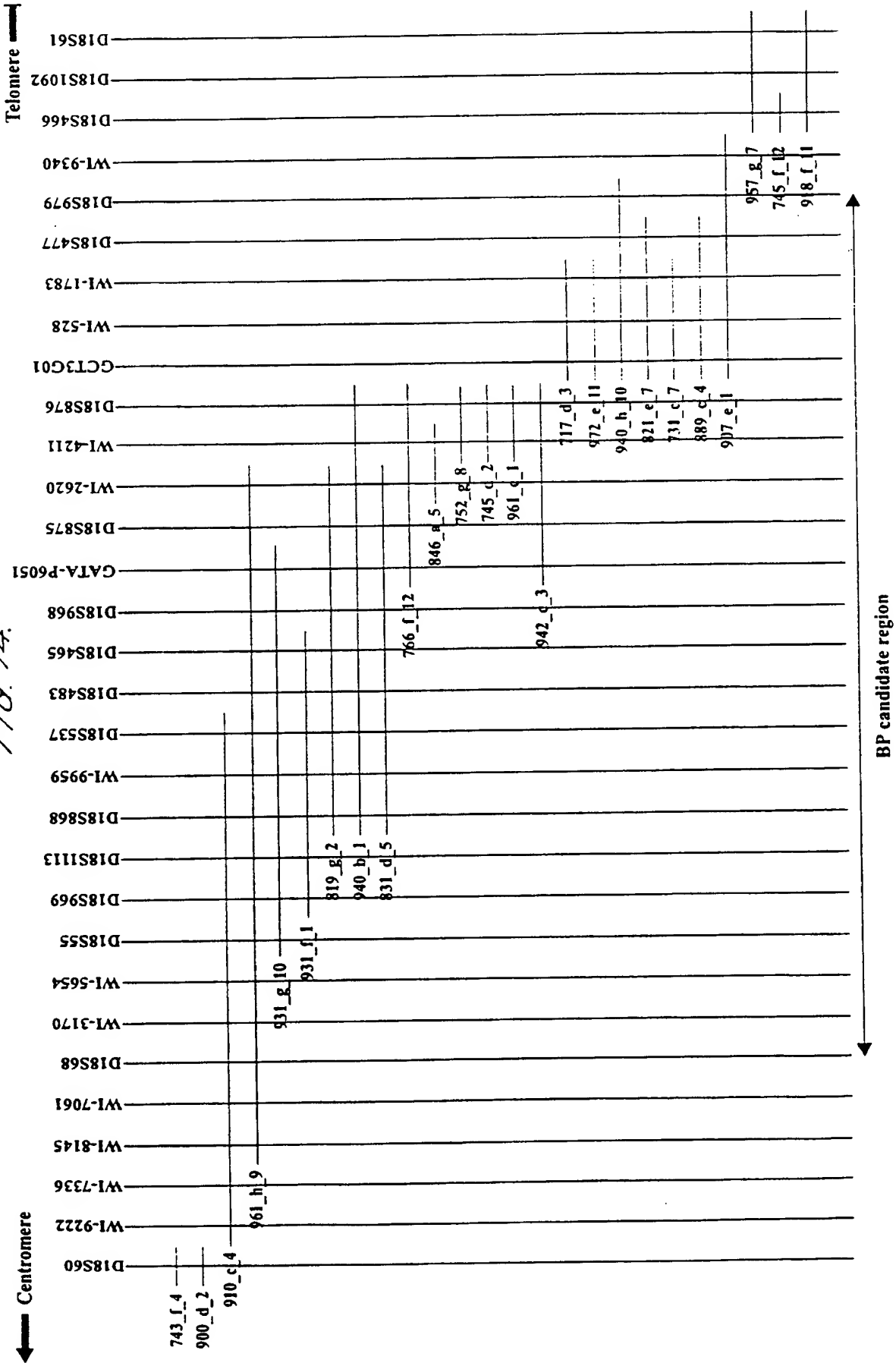
Telomere
D18S61
D18S1092
D18S466
WI-9340
D18S979
D18S477
WI-1783
WI-528
GCT3G01
D18S876
WI-4211
WI-2620
D18S875
GATA-P6051
D18S968
D18S465
D18S483
D18S537
WI-9959
D18S868
D18S1113
D18S969
D18S55
WI-5654
WI-3170
D18S68
WI-7061
WI-8145
WI-7336
WI-9222
D18S60
Centromere
743_f_4
900_d_2
910_c_4
961_h_9
931_g_10
931_f_1
819_g_2
940_b_1
831_d_5
766_f_12
846_g_5
752_g_8
745_c_2
961_c_1
942_c_3
717_d_3
972_e_11
940_h_10
821_e_7
731_c_7
889_c_4
917_e_1
957_g_7
745_f_12
918_f_11
BP candidate region

FIG. 1A

FIG. 15a.

GCCAACAAACAAAATGAAATAAGACCTGGGATGTATTTTTTGGCCAAGGCAATTAGAAA
ATGATTAGTATCCTTATCAGGAGCAATTTCAGAGAATGTTTGGGTGGACGTCTAACTACA
GTGGAGTCAAACGTGAATCAACGGTGAAAAAAGGACAATAGCCAATGTGTACACTTTTT
ATAAAAACCACCCTCCAAGGACCAGGCACTGGCCCTCTCTCCGGTGCCCACAGACATC
CACACAGGCCCAAAGAATCAGGGATTGCACAAGCCAGAGCAATCGAACGGTTCTGAGT
CATCTGCCGGAAGCCTTGCCCTCAATCAAGGCGGACGTGAAGCATCTACAAAGGAGGA
*ATAGTCAAAG***CAGCAG**CGGCGGCGGCGGCGGCGG**CAGCAGCAGCAGCAGCAG**G
AGGTGGGGGCCTCTGCCAGGTACCGGGCGGGGCAGGCACGGAGGTGCCCAGGTT
CCCGCGGAGGCCACCTCTTCCCTGGAGTGCGTGAGAGAGGGGAAGGGAGGAAGG
CCAGAGCAGGAATCAGAGCGAGGCAAAGGCGGGCAGGAACTAXGAGAATGACS
GCGGGAGGCGGCCGGGAAAGAAAXTCTCGGGGCTGTGGGGGTCXCCCTGGCACC
AGCCGGGGTCCCAAGCCCCACCGCGAGACCCCGCGA

FIG. 15b.

5'-ATCGAACGGTTCTGAGTCATCT
5'-CGCTCTGATTCCTGCTCTG

FIG. 16a.

TTCAGTAGAAGGAAGCACAGCAAATTTGCCTTTATAGAGATTCAATTCTTGGTGCTTGG
GCCAAAGAATAAGAATTACATTAAGCAGGCCGGGCACGGTGGCTCACACCTGTAAAAC
CAGAACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAGATCGAGACCATCC
TGGACAACATAGTGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCGGGCATGGTG
GTGCATGCCTGTAATCCCAGCTACTCAGGAGGCGGAGGCAGGAGAATCCCTTGAACCA
GGGAGTTGGAGGTTGCAGTGAGCCGAGATCACGCCACAGCACTCTAGCCTGGCGACA
*GAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAATTACATTAAG***CAGCAGC**
AGCAGCAGTGASAGAGGGAAKAATGAAAGAAGAAATTTCTAGAATAAGATTGA
TCTCCAGCACCATGCCAATCATGGACTGGATACAATTCATGCATATCTTTTGTGA
GAGAGGTGAGAGATGTGAATCCTTTCTCATT

FIG. 16b.

5'-AGAAGGAAGCACAGCAAATTTG
5'-GCATGGTGCTGGAGATCAAT

FIG. 17a.

TGGGAGTTAAAGCAGACATTCGGCTTTNGTGTTGCCAGAGTTCTAACATAAGTTCTTTTT
CATCTGGGCAGGCNGATGTTCCTTCCATCTTNGAAGNACNGTCCTTTTCATTTTTTTTAT
TTNGCTTTTGGSKTTTATCTTCTTAGACGTCTTCAGGAGTTKGATTGTAGKGTAAGGCAG
ATTTAGTTGACTGGGCTTTGTTTCTGGAAAATTTTAAAGGGGCAAGTCCTGGGCTGCAT
ATTCTTACTCTGGGGGCTTAGTACTGGCCCCTAAATTTGTTCTCTGGCTCCTCAAGGTT
AGAAATCTGCTGGCTGGAGGGGCTGAGATGTTCCTTGACTGCTGGCCAGAACATTCCG
CCGGGGGGTGGCAACCGAAGTGTTTCTTTGGGCAATGGCAGCAGAATTCATGATTGTT
*TTCATGTRC***CAGCAGCAG**TGGCAGCGCAKTGAGTTGCATGATTGTTGGCTGGGGC
TGAGTGCTGGCASGCACTGGAGTGTTTGGCTTCCAGTAGAAATTCACAGCAGTAG
TAGTGGTGGCATGGGAAGGAGGGCAGYGGTGGCATGGGGAGGACCCCCC

FIG. 17b.

5'-GGCTGAGATGTTCCTTGACTGC
5'- CCTTCCCATGCCACCACTACTA

FIG. 18a.

TGTAATTCCCAGCAATTTGGGGAGCCCAAGGCGGGCAGATTCATGAGTTCGGGAAGAT
TCGAGACCNTTCCTGGCTAAACACGGGGGAAACCCCCNTTTTTACTAAAAAATACCAAAA
AATTAACCTGGGCGTGGTGGCGGGCCCCAGCTANTCCGGAGGCTGAGGCAGGAGAAT
GGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCCCGCTACTGCACTCCA
GCCTGGGCAATAGAGGGAGACTCCGTCTCAAAAAAAAAAAAAAAATAAATAATAATAAAA
AAAATAACAATAATAATACTAATAATTGCTTGATATTTTACAAAAGCAAAAGGAAAAGAAG
ACTAGGCAAGAAAAAAAAAACCTCCTTAGATGGTAGAACTCAGGTTTAAAATTAAAACTT
ATTCTGGTGTCAGSCTAGTTGTATATTTTGACCTCTTTAAATGCTCTGAACTATGATATGG
*AGTAACAGCGATG***CTGCTGCTGCTGCTGCTGCTGCTG**ATGGTGGTGGTGTTTTA
ATATCGAATAAAAGTTGTGGAAACTAAATTTCATTTCTGCCAATTAACTAAGATT
GCAAAGTTAAACATCT

FIG. 18b.

5'-TTTGCAATCTTAGTTAATTGGC
5'-GAACTATGATATGGAGTAACAGCG